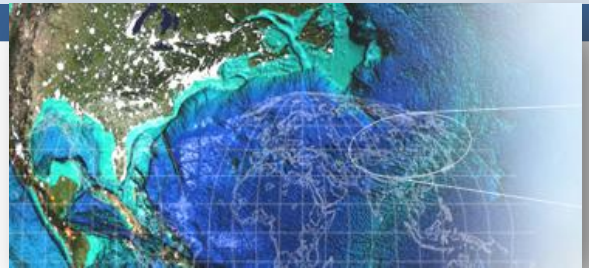
ESIP Federation Interoperability Outreach Package

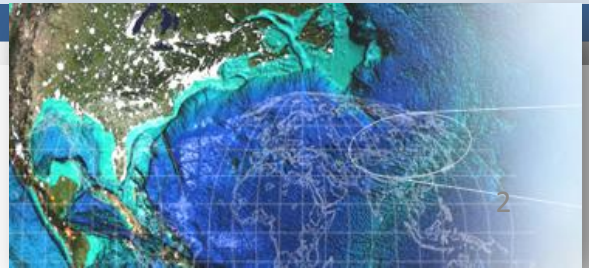
December 7, 2010
DRAFT

Federation of Earth Science Information Partners
MAKING DATA MATTER



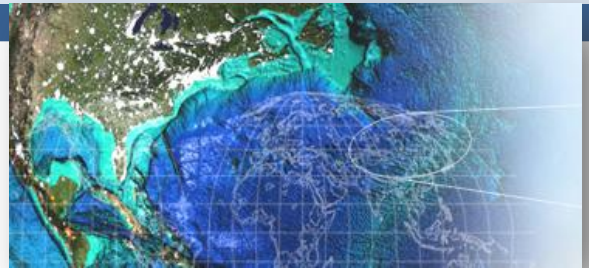
Outline

- Background
- Interoperability by Functionality
  - Introduction
  - Data Access
    - Open Geospatial Consortium (OGC)
    - Open-source Project for a Network Data Access Protocol (OPeNDAP)
  - Catalog Search and Discovery
  - Usage and Formats
  - Semantic Web
- Case Studies
- Computing Platforms
- References



Federation of Earth Science Information Partners
MAKING DATA MATTER

Background



Federation of Earth Science Information Partners
MAKING DATA MATTER

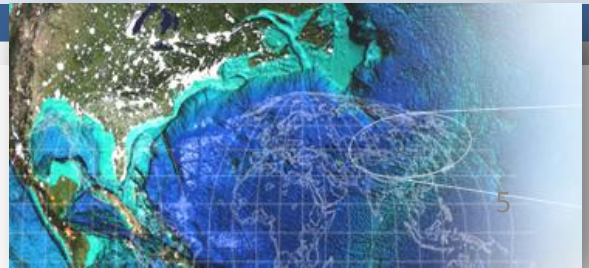
About the ESIP and IT&I Cluster

- The ESIP Federation is a consortium of more than 110 organizations that collect, interpret, and develop applications for Earth observation information.
- Partners include NASA and NOAA as strategic funding partners (Type IV), USGS data centers, and research universities, and many other organizations involved in Earth science.
- ESIP's Information Technology and Interoperability cluster was created, in part, "to ensure that data, information and services can be readily discovered, exchanged and integrated through the use of interoperability standards and protocols." http://wiki.esipfed.org/index.php/IT%26I_Chair

Federation of Earth Science Information Partners
MAKING DATA MATTER

Outreach Goal

- The ESIP Federation initiative coupled with community standards, such as those of the Open Geospatial Consortium (OGC) and existing legacy systems, will allow different scientific communities to work together to achieve a bigger vision.



Federation of Earth Science Information Partners
MAKING DATA MATTER

About Interoperability

- Interoperability is a feature of useful network of sharing information that provides value to its participants allowing the data providers to compete to be part of these system of systems.
- The IEEE (Institute of Electrical and Electronics Engineers) defines interoperability as: “the ability of two or more systems or components to exchange information and to use the information that has been exchanged.”

http://www.newworldencyclopedia.org/entry/Interoperability

Federation of Earth Science Information Partners
MAKING DATA MATTER

Types of Interoperability

Syntactic interoperability

- "If two or more systems are capable of communicating and exchanging data, they are exhibiting syntactic interoperability. Specified data formats, communication protocols and the like are fundamental. In general, XML or SQL standards provide syntactic interoperability. This is also true for lower-level data formats, such as ensuring alphabetical characters are stored in ASCII format in both of the communicating systems."
- Syntactical interoperability is a necessary condition for further interoperability.
- http://en.wikipedia.org/wiki/Interoperability

Semantic interoperability

- "Beyond the ability of two or more computer systems to exchange information, semantic interoperability is the ability to automatically interpret the information exchanged meaningfully and accurately in order to produce useful results as defined by the end users of both systems. To achieve semantic interoperability, both sides must defer to a common information exchange reference model. The content of the information exchange requests are unambiguously defined: what is sent is the same as what is understood."
- http://en.wikipedia.org/wiki/Interoperability

Approaches to Interoperability

There are two approaches:

- Luxury of building a new system with high costs
- Utilizing the pre-existing legacy system and bringing up to speed

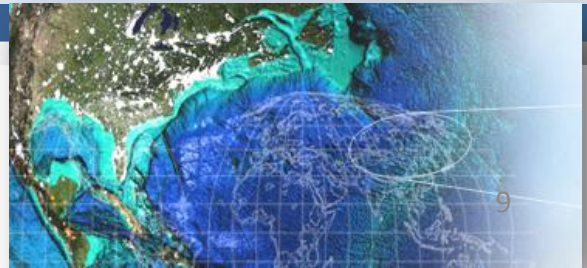
Regardless of which path is selected, both require:

- A service-oriented approach that underlies much of today's World Wide Web
  - New systems shall use a service-oriented approach to making their data available to distributed communities of scientists
- transaction-oriented and conceptual aspects of Web 2.0

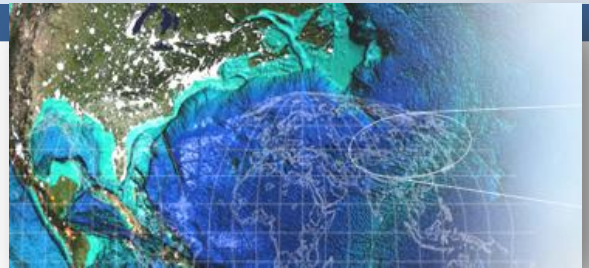
Federation of Earth Science Information Partners
MAKING DATA MATTER

Interoperability Standards

- Open Geospatial Consortium (OGC) Standards
- ESIP Federation Open Search
- Open-source Project for a Network Data Access Protocol (OPeNDAP)
- Metadata
- File Formats



Interoperability by Functionality

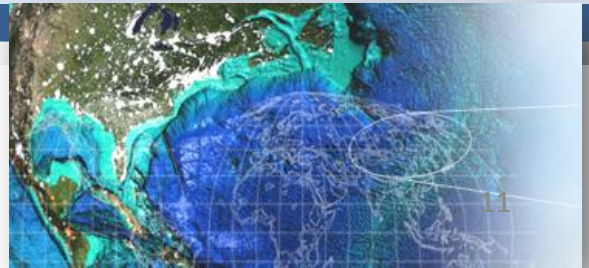


Federation of Earth Science Information Partners
MAKING DATA MATTER

Interoperability by Functionality (1/2)

This Outreach package focuses on interoperability based on three main functionalities:

- Data Access
- Catalog Search and Discovery
- Usage and File Formats



Interoperability by Functionality (2/2)

Data Access

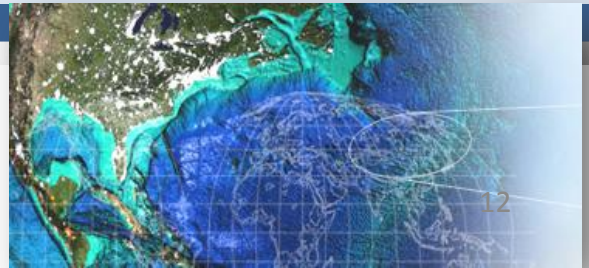
- OGC Standards
  - WMS, WCS, WFS
  - SWE, SSW, GSW
- OPeNDAP

Catalog Search And Discovery

- ESIP Open Search
- CS-W
- ebRIM
- GEOSS Search & Discovery
- Federated Search

Usage and File Formats

- File Formats
  - FDGC
  - NetCDF and CF
  - ISO 19115 and 19119
  - NASA ECS and GCMD DIF



Federation of Earth Science Information Partners
MAKING DATA MATTER

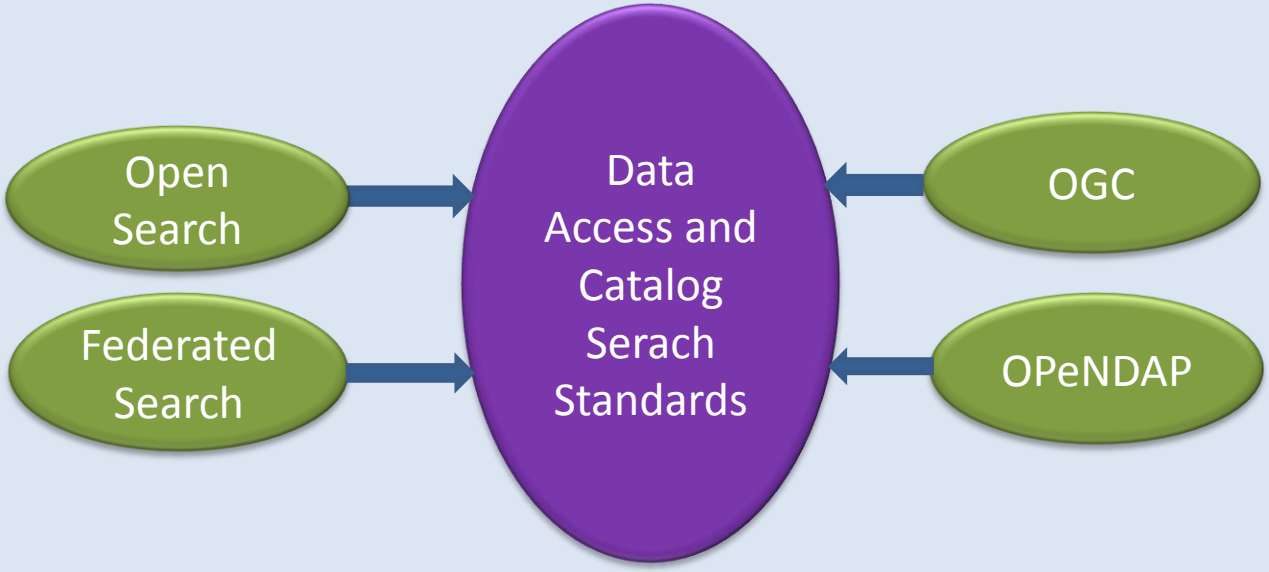
OGC Participants/Communities

There are three important groups of participants:

1. Data providers who should be able to download an API to implement OGC web services to distribute their data.
2. The community of users who can utilize a web-based visualization tool or client tool that is capable of communicating with data providers in a form such as REST, SOAP, or XML.
3. Developers who implement, test, and integrate these systems/software and provide them to the user community. This developer community is formed by a group of experts in software engineering, standards, and database managers.

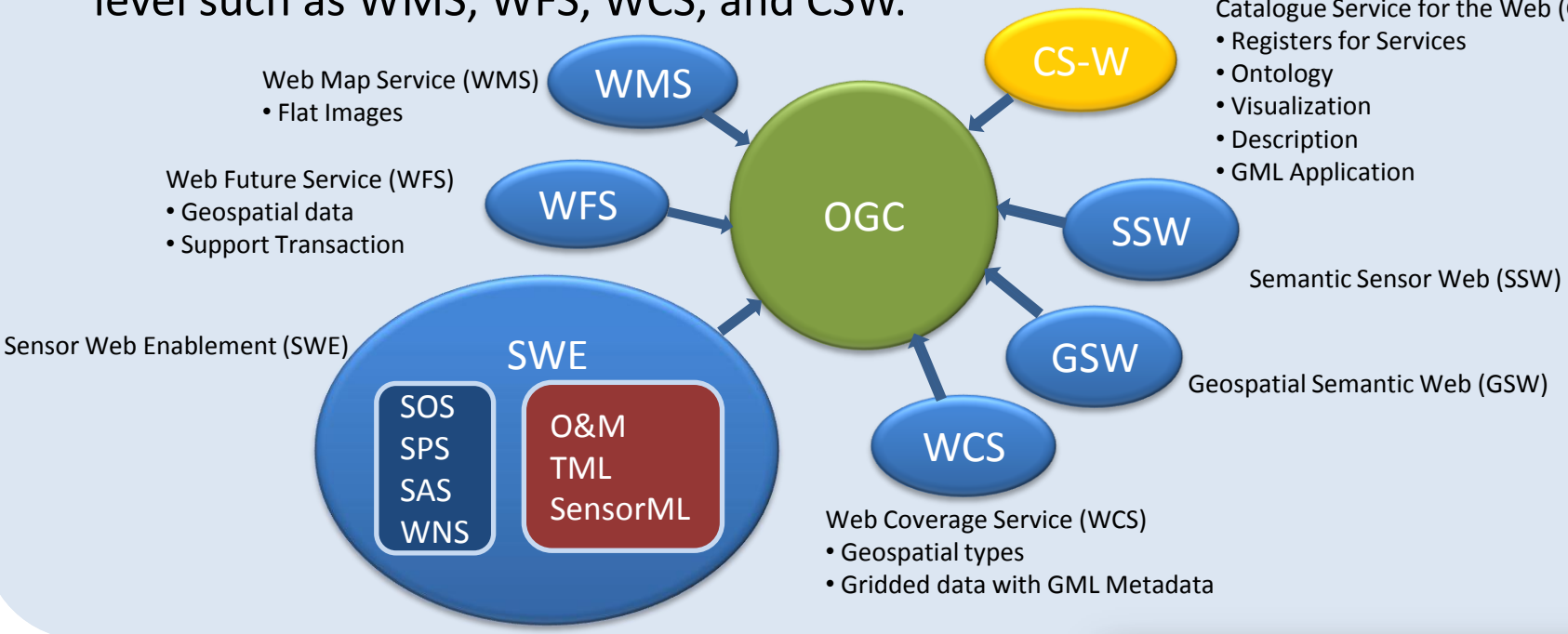
Interoperability Standards

- Interoperability standards include ones developed by the Open Geospatial Consortium (OGC), Federated Search, and OPeNDAP.



OGC Standards

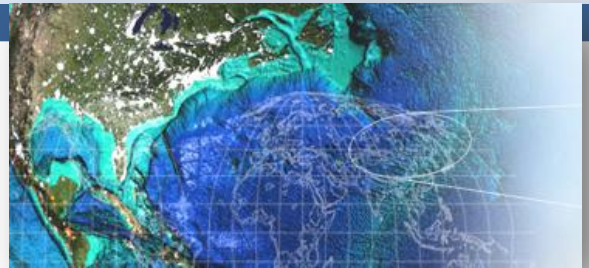
- The Open Geospatial Consortium (OGC) has, via their standards process, developed protocol standards implemented operationally at various level such as WMS, WFS, WCS, and CSW.



Federation of Earth Science Information Partners
MAKING DATA MATTER

Data Access – OGC

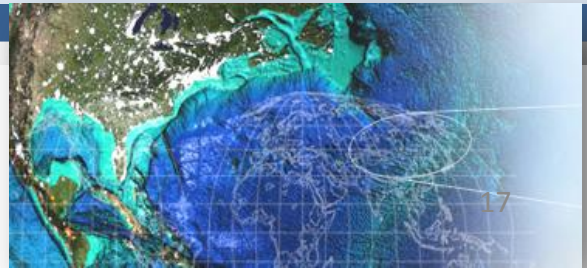
Open Geospatial Consortium



Federation of Earth Science Information Partners
MAKING DATA MATTER

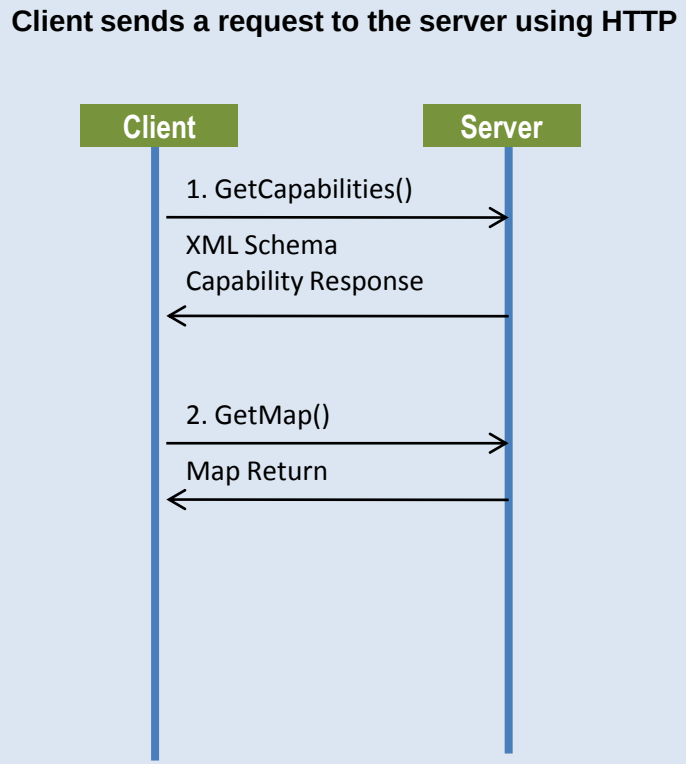
Some OGC Standards

- Web Map Service (WMS)
- Web Feature Service (WFS)
- Web Coverage Service (WCS)
- Sensor Web Enablement (SWE)
- SensorML



Web Map Service (WMS) – Maps

- OGC WMS standards serve visualizations of geographical data in “flat” map-like layered images.
- HTTP queries to a WMS server return pre-rendered images from different sources.
- The returned images from different sources can be layered on top of each other.

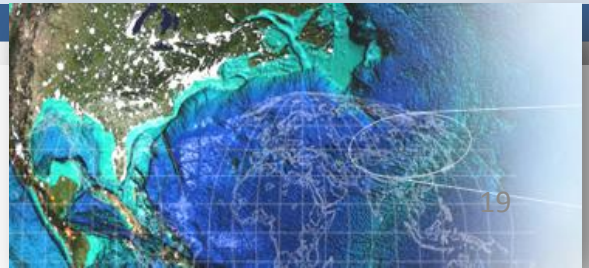


Federation of Earth Science Information Partners
MAKING DATA MATTER

WMS Examples (1/3)

- The following links are two example for GetCapabilities Request:
- Northern Hemisphere: http://nsidc.org/cgi-bin/atlas_north?service=WMS&request=GetCapabilities&version=1.1.1
- Southern Hemisphere: http://nsidc.org/cgi-bin/atlas_south?service=WMS&request=GetCapabilities&version=1.1.1

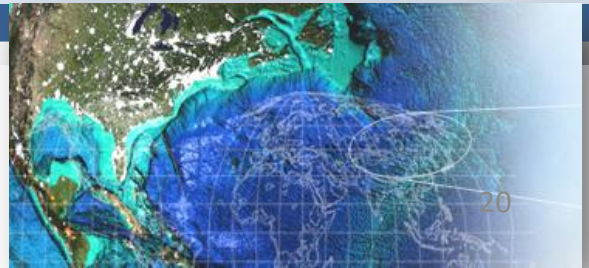
http://nsidc.org/data/atlas/ogc_services.html



WMS Examples (2/3)

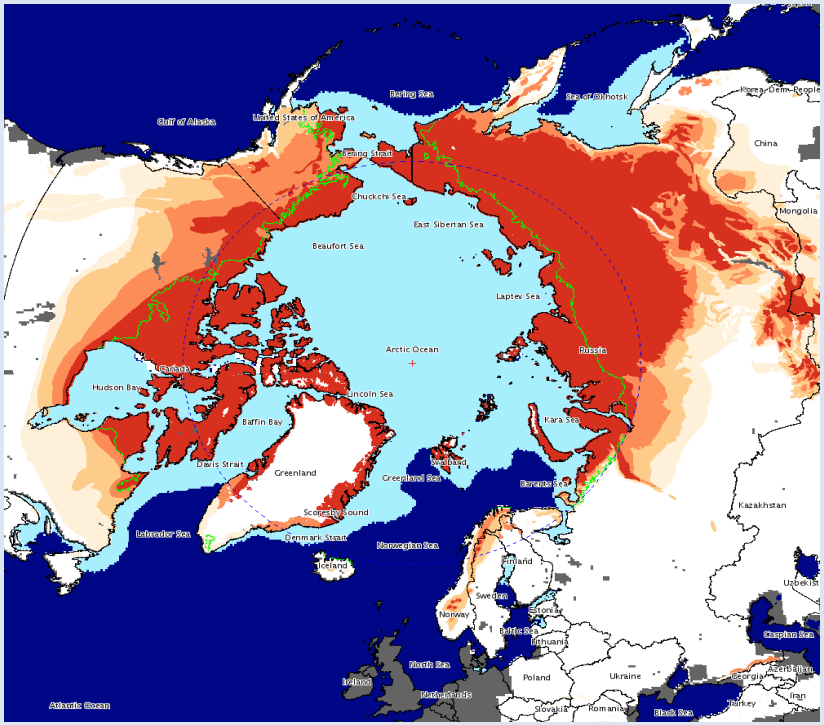
- GetMap Northern Hemisphere request using the known parameters from GetCapabilities:

http://nsidc.org/cgi-bin/atlas_north?
service=WMS&
version =1.1.1&
request =GetMap&
srs=EPSG:32661&
format=image/gif&
width=1000&
height=1000&
bbox=-2700000,-2700000,6700000,6700000&
layers=sea_ice_extent_01,land,snow_extent_01,permafrost_extent,country_borders,treeline,north_pole_geographic,arctic_circle,country_labels,geographic_features_sea



WMS Examples (3/3)

- GetMap Response from Server:



Federation of Earth Science Information Partners
MAKING DATA MATTER

Web Feature Service (WFS) – Features

WFS Features

Publishes feature-level geospatial data to the web and supports INSERT, UPDATE, DELETE, LOCK, QUERY, and DISCOVERY operations on these data using HTTP

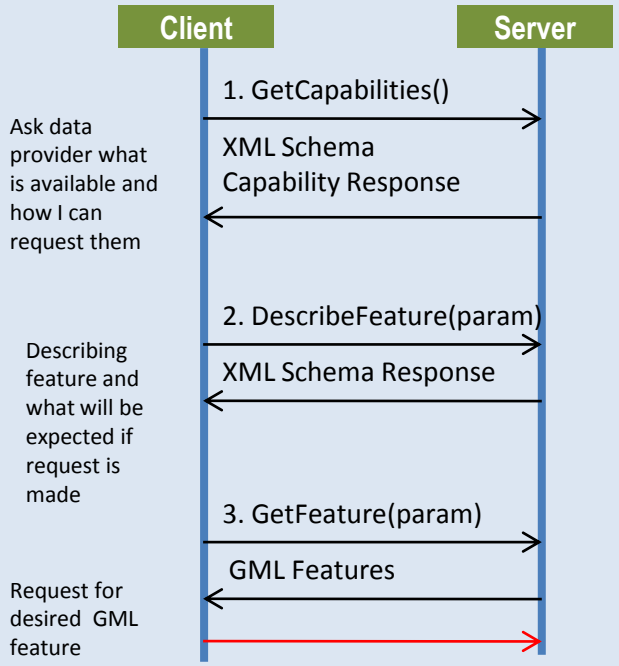
WFS Provide three main Requests:

- GetCapabilities: Describes its which feature types it can service and what operations are supported on each feature type.
- DescribeFeatureType: Describes the structure of any feature type it can service
- GetFeature: Allows retrieval of feature instances based on client specified request to constrain the query spatially and non-spatially.

WFS Optional Requests:

- GetGmlObject: Provides retrieval of element instances by traversing XLinks that refer to their XML IDs.
- Transaction : Allows create, update, and delete operations on geographic features
- LockFeature: Provides lock request on one or more instances of a feature type for the duration of a transaction to ensure support of serializable transactions

Client sends a request to the server using HTTP



Federation of Earth Science Information Partners
MAKING DATA MATTER

WFS Benefits

- Instead of returning an image as WMS does, WFS provides detailed information about specific geospatial features of the underlying data, at geometry and attribute levels.
- Unlike the WMS getFeatureInfo request that only returns information about the feature, WFS provides the geometry itself.
- Whereas WMS offers imaging services, WFS provides geographical features as the “source code” behind the map.
- WFS has optional features that allow inserting and modifying geospatial objects such as weather observation readings.

WFS Examples (1/3)

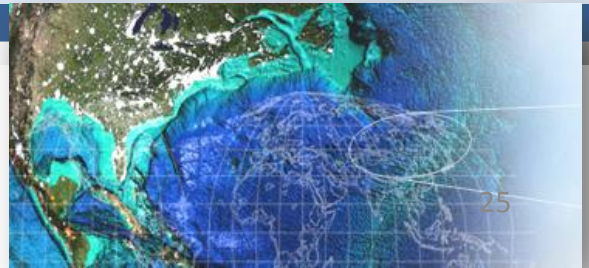
- The following links are two example for GetCapabilities Request:
- Northern Hemisphere: http://nsidc.org/cgi-bin/atlas_north?service=WFS&request=GetCapabilities&version=1.1.0
- Southern Hemisphere: http://nsidc.org/cgi-bin/atlas_south?service=WFS&request=GetCapabilities&version=1.1.0

http://nsidc.org/data/atlas/ogc_services.html

WFS Examples (2/3)

- GetFeature Request: Elevation contours for the Greenland ice sheet:

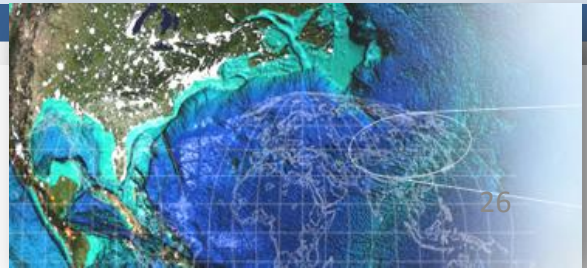
  http://nsidc.org/cgi-bin/atlas_north?
  service=WFS&
  version=1.1.0&
  request=GetFeature&
  typename=greenland_elevation_contours



WFS Examples (3/3)

- WFS GetFeature XML Response from Server:

<wfs:FeatureCollection xsi:schemaLocation="http://mapserver.gis.umn.edu/mapserver http://nsidc.org/cgi-bin/atlas_north?SERVICE=WFS&VERSION=1.1.0&REQUEST=DescribeFeatureType&TYPENAME=greenland_elevation_contours&OUTPUTFORMAT=text/xml;
subtype=gml/3.1.1 http://www.opengis.net/wfs http://schemas.opengis.net/wfs/1.1.0/wfs.xsd">
<gml:boundedBy><gml:Envelope srsName="EPSG:32661">
<gml:lowerCorner>-406699.951843 -456661.863393</gml:lowerCorner>
<gml:upperCorner>1677349.428068 1589388.039689</gml:upperCorner>
</gml:Envelope>
</gml:boundedBy>
<gml:featureMember>
<ms:greenland_elevation_contours><gml:boundedBy>
<gml:Envelope srsName="EPSG:32661“>
<gml:lowerCorner>-327189.501776 -456661.863393</gml:lowerCorner>
…….
……
</gml:featureMember>
</wfs:FeatureCollection>

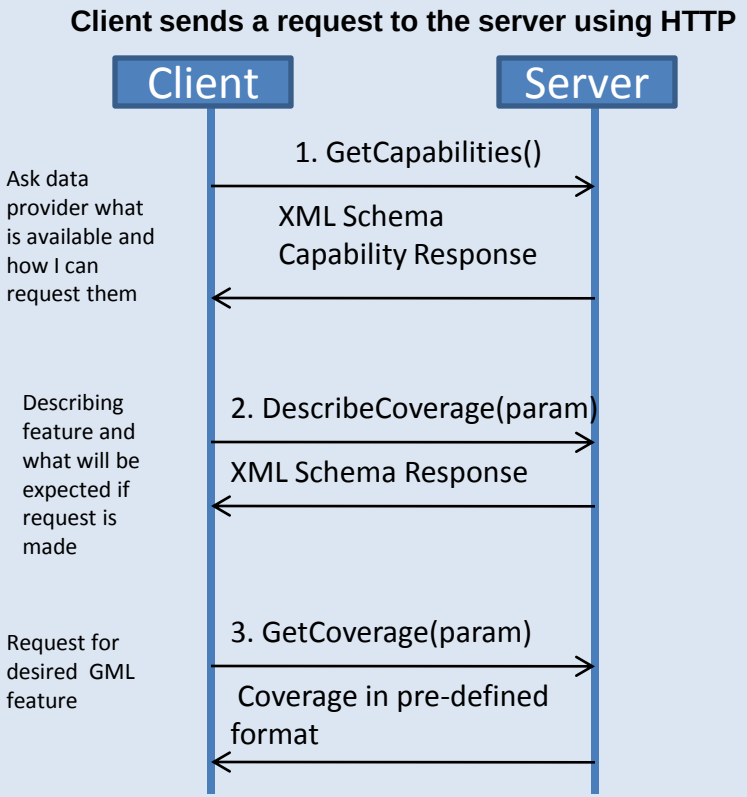


Federation of Earth Science Information Partners
MAKING DATA MATTER

Web Coverage Service (WCS) – Gridded data

Web Coverage Service (WCS)

- **- Gridded data with GML metadata**
- Enables interoperable access to detailed and rich geospatial data as “coverages”
- Supports Binary files: GeoTiff, HDF-EOS, CF-NetCDF, JPEG
- Provides three main Requests:
  - GeCapability: Retrieves a list of the server’s data, available Services, WCS operations, and parameters
  - DescribeCoverage: Retrieves an XML document that fully describes the requested coverages
  - GetCoverage: Returns a coverage in a well-known format with several extensions to support the retrieval of coverages



Federation of Earth Science Information Partners
MAKING DATA MATTER

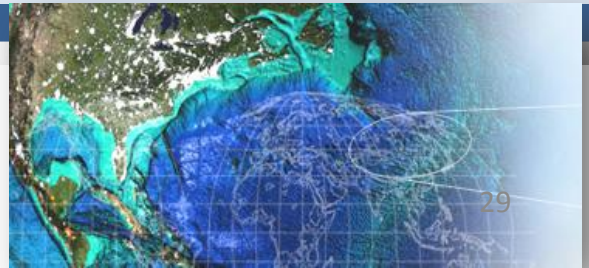
WCS Benefits

- Like WMS, the WCS standard returns similar formats
- In addition, WCS is capable of providing:
  - valuable metadata and more formats
  - Multi-valued coverages
  - more precise queries against multi-dimensional backend formats
  - raster data in form of ‘source code’ with detail descriptions of the map unlike WMS raw static image
  - coverages that relate a multidimensional spatio-temporal domain to a range of properties unlike WFS which returns discrete geospatial features
- http://docs.geoserver.org/latest/en/user/services/wcs/reference.html

WCS Examples (1/3)

- The following links are two example for GetCapabilities Request:
- Northern Hemisphere: http://nsidc.org/cgi-bin/atlas_north?service=WCS&request=GetCapabilities&version=1.1.1
- Southern Hemisphere: http://nsidc.org/cgi-bin/atlas_south?service=WCS&request=GetCapabilities&version=1.1.1

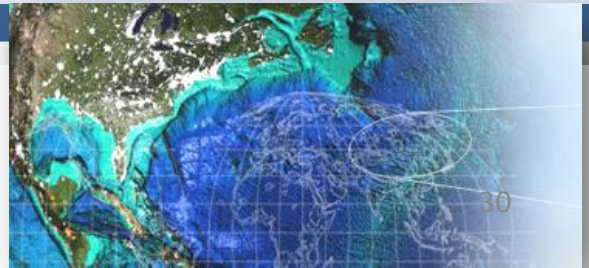
http://nsidc.org/data/atlas/ogc_services.html



WCS Examples (2/3)

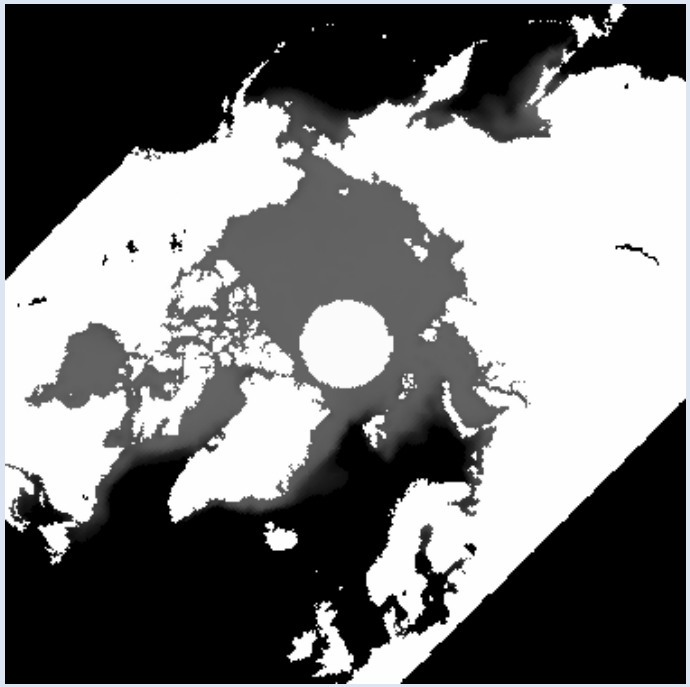
- GetCoverage : After requesting GetCapabilities and DescribeCoverage, the WCS Client can place a GetCoverage Request. The below is a GetCoverage REST request for A GeoTIFF of sea ice concentration at 25-km resolution in a polar stereographic projection focused on the Arctic (138 KB):

  http://nsidc.org/cgi-bin/atlas_north?
  service=WCS&
  version=1.1.1&
  request =GetCoverage&
  crs=EPSG:32661&
  format=GeoTIFF&
  resx=25000&
  resy=25000&
  bbox=-2700000,-2700000,6700000,6700000&
  coverage=sea_ice_concentration_01



WMS Examples (3/3)

- GetCoverage Response from Server:



Sensor Web Enablement (SWE)

- Enables interoperability among different sensor types, different systems, and different groups:
  - Remote sensor vs. in-situ
  - Emergency system, utility, defense, science, and intelligence
  - Government vs. commercial
- Supports any sensor
- Supports semantic web on-line dictionary and ontology
- Task sensors
- Subscribe to sensor published alerts

SWE Goals

OGC SWE goal is to discover:

- Sensor systems
- Observations and processes
- Sensor capabilities and quality of measurements
- Sensor parameters that allow software to process and geolocate observations
- Methods to retrieve real-time or time series observations and coverages
- http://portal.opengeospatial.org/files/index.php?artifact_id=14140

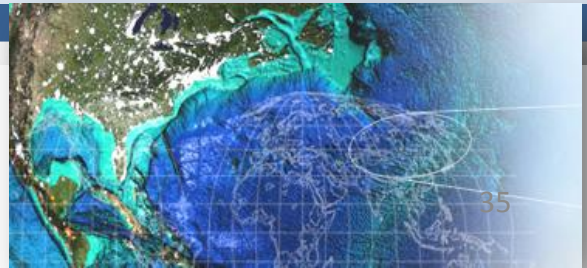
Sensor Web Enablement Models

SWE is a Service Oriented Architecture (SOA) open standard providing:

- Four standard web services interface models
  - Sensor Observation Service (SOS)
  - Sensor Alert Service (SAS)
  - Sensor Planning Service (SPS)
  - Web Notification Service (WNS)
- Three XML encoding information models
  - Sensor Model Language (SensorML)
  - Observations and Measurements (O&M)
  - Transducer Markup Language (TML)

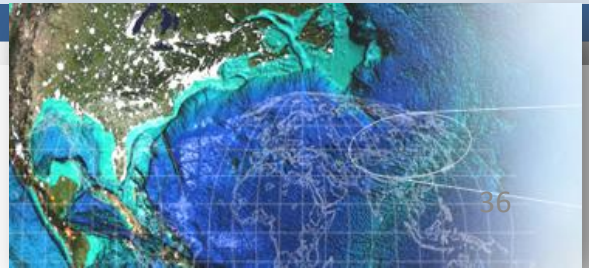
SWE Standard Web Services Interface Models

- Sensor Observation Service (SOS)
- Sensor Alert Service (SAS)
- Sensor Planning Service (SPS)
- Web Notification Service (WNS)

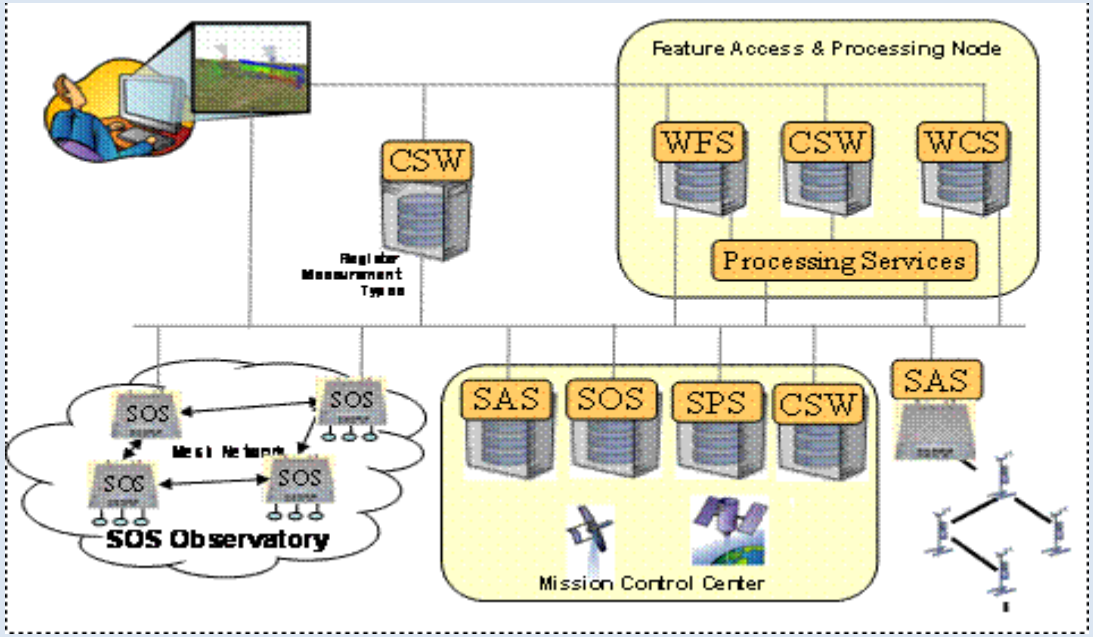


SWE XML Encoding Information Models

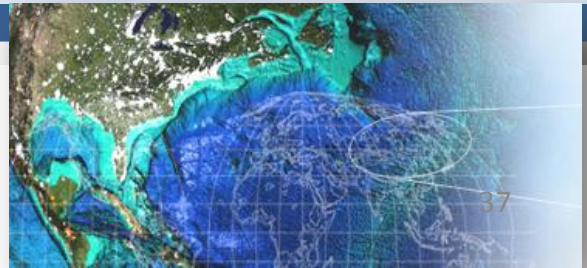
- Sensor Model Language (SensorML)
- Observations and Measurements (O&M)
- Transducer Markup Language (TML)



SWE Framework



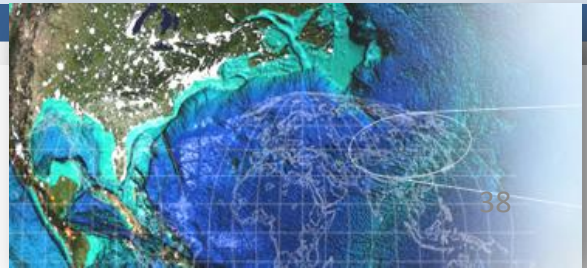
http://www.opengeospatial.org/standards/orm



Federation of Earth Science Information Partners
MAKING DATA MATTER

SWE Visualization Clients

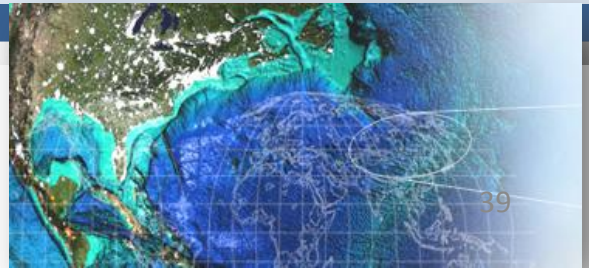
- http://code.google.com/p/space-time-toolkit/
- http://www.crisisgrid.org/html/ogc-swe.html



Federation of Earth Science Information Partners
MAKING DATA MATTER

SWE Examples

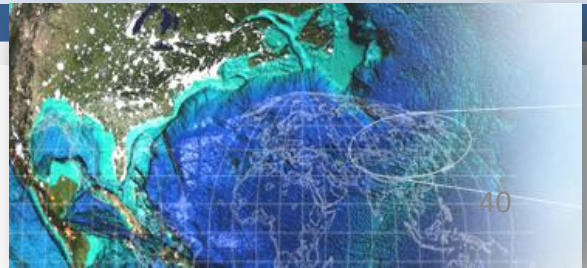
- To Be Completed
- Some SWE-enabled software is listed at http://www.ogcnetwork.net/SWE_Software
- Client (with download links)
- Server (with download links)
- Picture examples



SensorML (1/3)

- “The OpenGIS® Sensor Model Language Encoding Standard (SensorML) specifies models and XML encoding that provide a framework within which the geometric, dynamic, and observational characteristics of sensors and sensor systems can be defined.”

http://www.opengeospatial.org/standards/sensorml



SensorML (2/3)

- The Objective of SensorML is to provide a complete description of an instrument's capabilities and give the information needed to process and geolocate the measured data.
- The information includes:
  - Sensor name, type, and identification numbers
  - Temporal, legal, or classification constraints of the description
  - A reference to the platform description
  - The sensor's coordinate reference system definition
  - The sensor's location
  - The response characteristics and information for geolocating samples
  - The sensor operator and tasking services
  - Textual metadata and history of the sensor
  - Textual metadata and history of the sensor description document itself
- http://www.sensorsmag.com/networking-communications/a-sensor-model-language-moving-sensor-data-internet-967

SensorML (3/3)

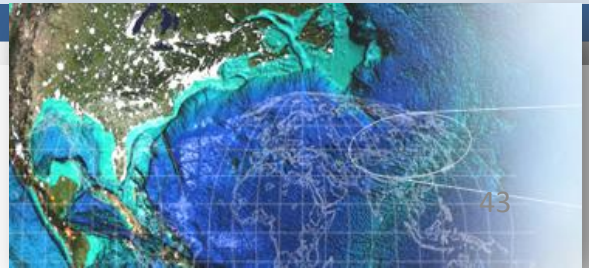
- By describing sensors using SensorML, anyone can put sensors or sensor data online for others to find and use.
- With automation of reading metadata, SensorML will enable development of software tools that automatically co-register different kinds of data with little human intervention.
- SensorML makes sensors become more intelligent and autonomous by assisting in onboard processing of data and communications among multiple sensors

http://www.sensorsmag.com/networking-communications/a-sensor-model-language-moving-sensor-data-internet-967

SensorML Examples

- Some SensorML software is listed at http://www.ogcnetwork.net/SWE_Software

- Client (with download links)
- Server (with download links)
- Picture examples



Federation of Earth Science Information Partners
MAKING DATA MATTER

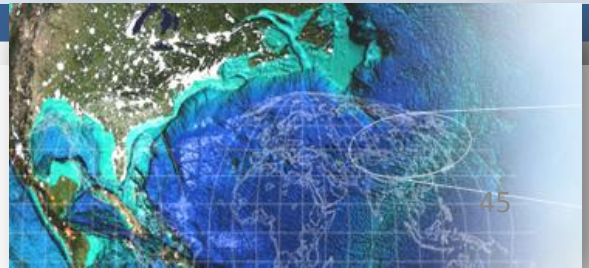
Data Access – OPeNDAP

Open-source Project for a Network Data Access Protocol

Federation of Earth Science Information Partners
MAKING DATA MATTER

OPeNDAP Data Access System

- Open-source Project for a Network Data Access Protocol (OPeNDAP) is a data transport and protocol, based on HTTP, used by Earth Scientists. See http://opendap.org/ for details.
- OPeNDAP specifies how to request a subset of a large dataset .
- It also specifies the over-the-network format for the response.
- The simplicity, flexibility, and efficiency of the protocol have led to its wide usage in the oceanographic and other scientific communities.
- There are many DAP compatible servers such as Hyrax, THREDDS, PyDAP, and ERDDAP.
- OPeNDAP servers manage large amounts of remotely sensed and model data, typically in grid formats such as NetCDF, GRIB, and HDF.



OPeNDAP Response Types

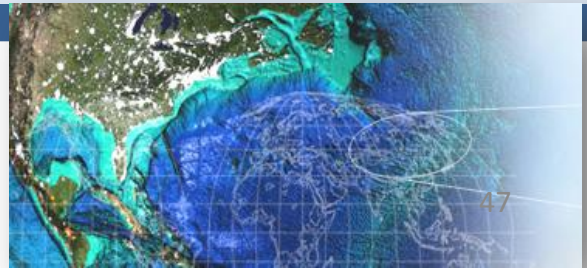
- OPeNDAP provides three basic data types.
- Request to Client OPeNDAP will result in following three different requests for data to OPeNDAP Server:
  - Request for the Data Descriptor Structure (DDS): is a data structure to describe datasets and subsets of those datasets – Shape of Data
    - Syntactic Metadata - Rigid
  - Request for the Data Attribute Structure (DAS): is a set of name-value pairs used to describe the data in a particular dataset
    - Semantic Metadata - Flexible
  - Request for actual Data (DODS): is data in a binary structure
- http://www.opendap.org/api/wc-html/writing_client_6.html

OPeNDAP Client/Server Architecture

- OPeNDAP Servers
  - CODAR
  - netCDF
  - HDF4/5
  - Matlab
  - DSP
  - Tables (JGOFS)
  - SQL (JDBC)
  - FITS
  - CDF
  - Flat Binary (FreeFrom)
  - CEDAR
  - General (ESML)
- See also http://opendap.org/faq/whatServers.html

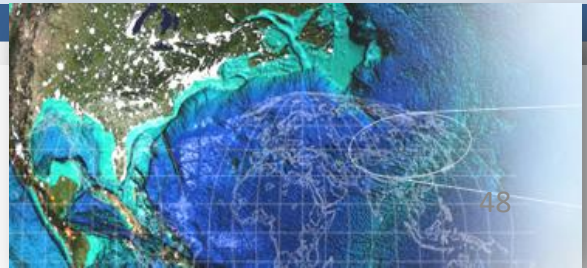
- OPeNDAP Clients
  - Ferret and GrADS (netCDF C)
  - IDV, VisAD, and ncBrowse (netCDF Java)
  - Matlab
  - IDL
  - Access
  - Excel
- See also http://www.opendap.org/faq/whatClients.html

- List of servers and clients given here from http://hdfeos.org/workshops/ws14/presentations/day1/OPeNDAP_tutorial_WS14.pptx



OPeNDAP Service at GSFC

- Currently, the GES DISC offers the datasets listed at http://disc.gsfc.nasa.gov/services/opendap/ through OPeNDAP.
- For example, AIRS data are made available through http://disc.sci.gsfc.nasa.gov/gesNews/opendap_AIRS_data_access



Federation of Earth Science Information Partners
MAKING DATA MATTER

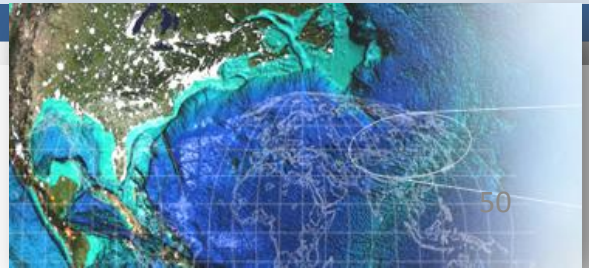
Uses of OPeNDAP

- OPeNDAP can be used to:
  - Make netCDF data files available over the Internet
  - Adapt existing software that use the netCDF API to read data served by an OPeNDAP data server
- In general, any program that uses netCDF can become a client in the OPeNDAP client-sever system.

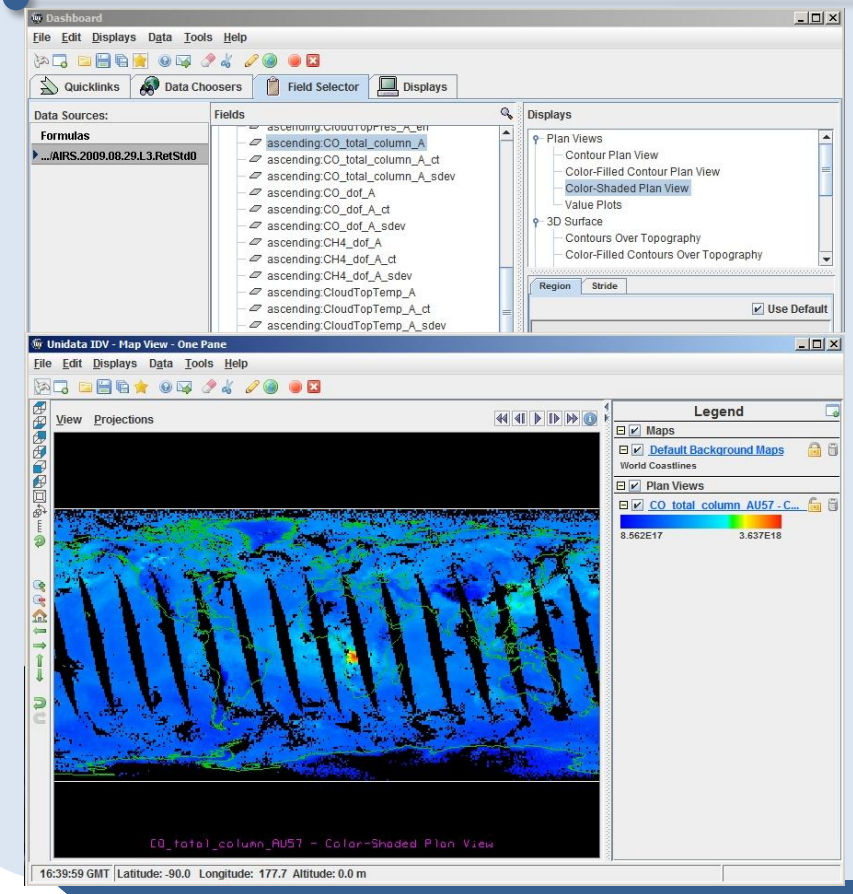
OPeNDAP Support of IDV, GrADS, and Panoply

- There are a variety of tools and file formats used with Earth science data, including IDV, GrADS, and Panoply.
- In general, IDV and Panoply have the easiest learning curves while GrADS is somewhat harder.
- OPeNDAP allows for remote access of data at a sub-file level, making the data easier to use.

http://wiki.esipfed.org/index.php/Making_Science_Data_Easier_to_Use_with_OPeNDAP



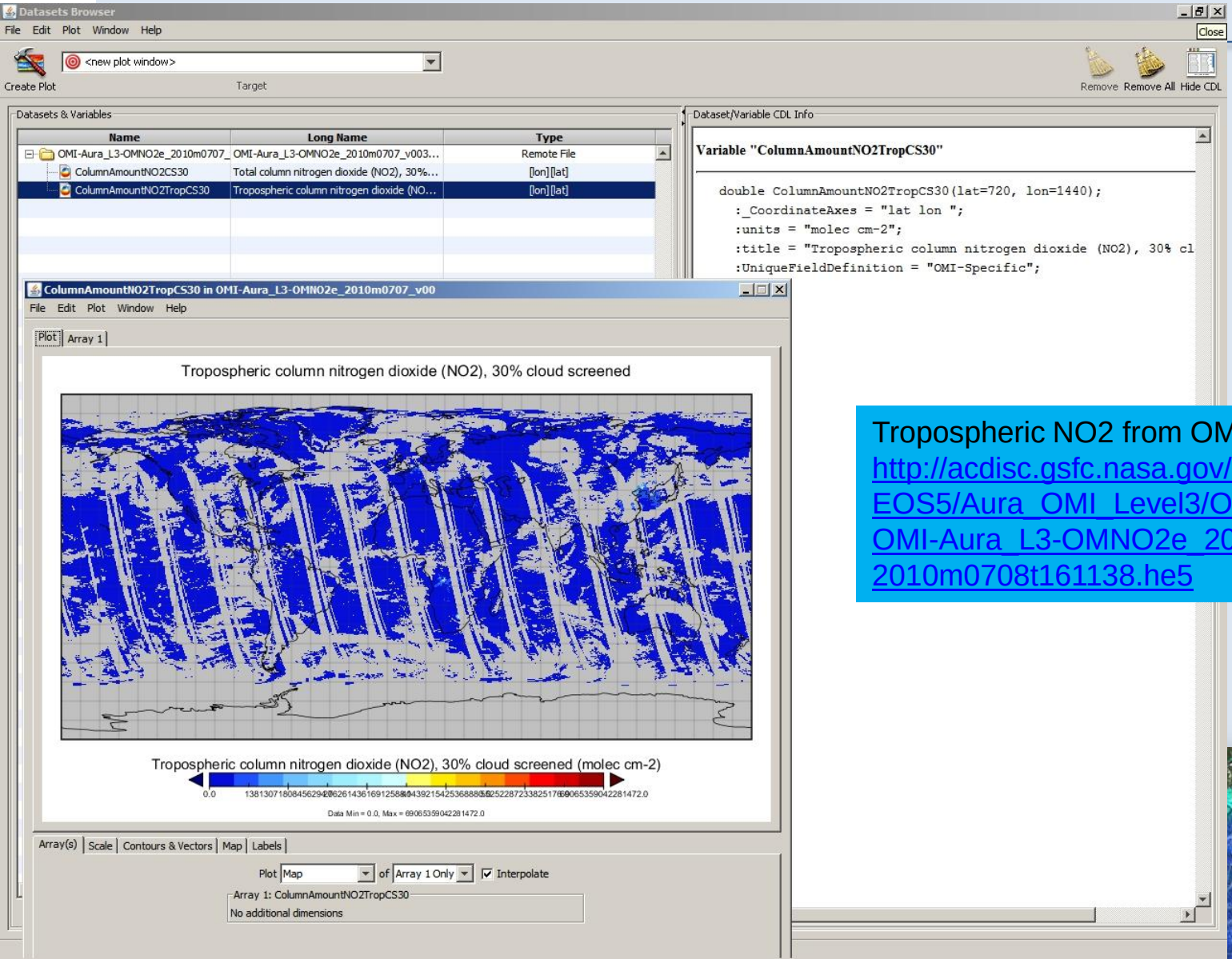
IDV Example



Carbon Monoxide Plume from California wildfires, seen in AIRS Daily Level 3 (AIRX3STD) on 29 Aug 2009: http://acdisc.gsfc.nasa.gov/opendap/Aqua_AIRS_Level3/AIRX3STD.005/2009/AIRS.2009.08.29.L3.RetStd001.v5.2.2.0.G09243131454.hdf

Federation of Earth Science Information Partners
MAKING DATA MATTER

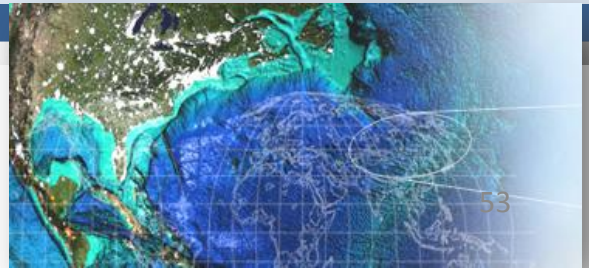
Panoply Example



Tropospheric NO2 from OMI for July 7, 2010: http://acdisc.gsfc.nasa.gov/opendap/HDF-EOS5/Aura_OMI_Level3/OMNO2e.003//2010/OMI-Aura_L3-OMNO2e_2010m0707_v003-2010m0708t161138.he5

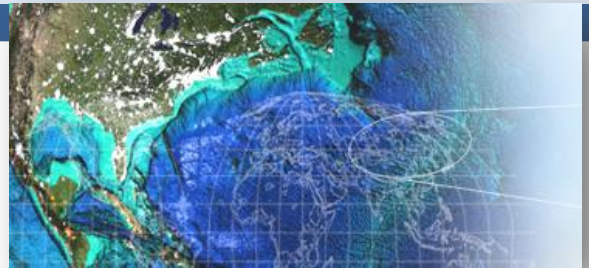
Other Examples

- Additional examples for Panoply and IDV, as well as examples for GrADS can be found at: http://wiki.esipfed.org/index.php/Making_Science_Data_Easier_to_Use_with_OPeNDAP
- This site also provides links to related tools.



Federation of Earth Science Information Partners
MAKING DATA MATTER

Catalog Search and Discovery

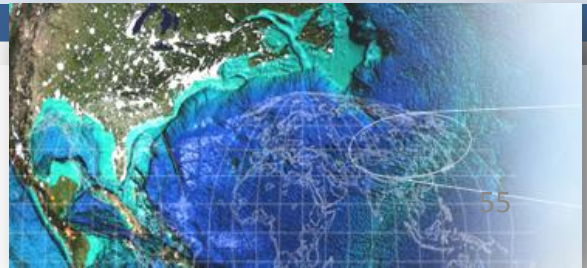


Federation of Earth Science Information Partners
MAKING DATA MATTER

Catalog Services – Web (CS-W)

- According to OGC specification, catalog services support the ability to publish and search:
  - Collections of metadata (representing resource characteristics)
  - Catalog services (required to support the discovery and binding to registered information resources)
  - Other related information objects

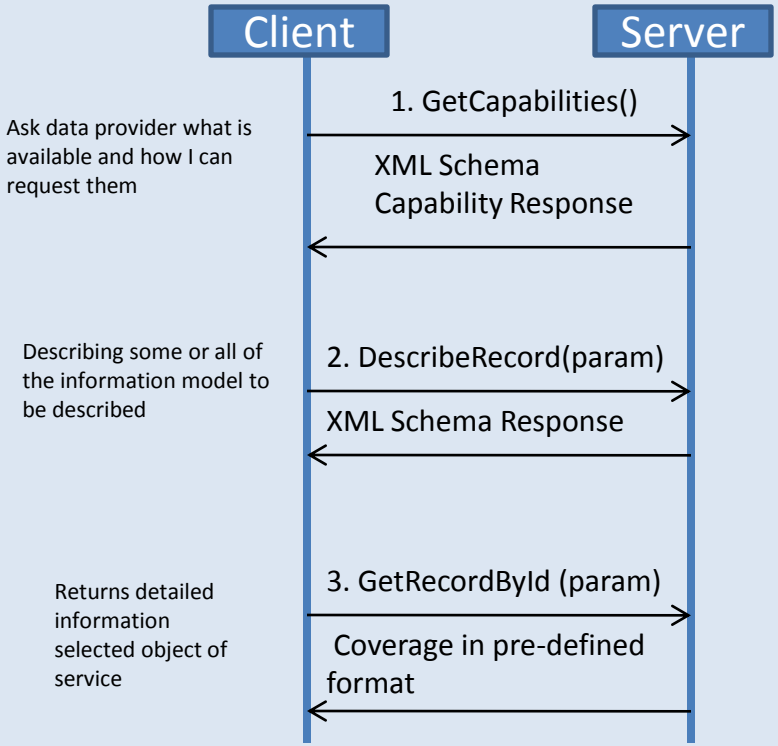
- http://www.opengeospatial.org/standards/cat



CS-W Abilities

CS-W Supports:

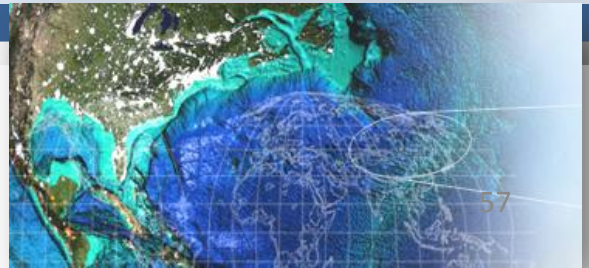
- GetCapabilities: Allows CS-W clients to retrieve service metadata from a server
- DescribeRecord: Allows a client to discover all data model supported by the catalogue service
- GetRecordById: Retrieves the default representation of catalogue records using their identifier.
- GetRecords
- http://eie.cos.gmu.edu/CSWClient/interface202.jsp



Federation of Earth Science Information Partners
MAKING DATA MATTER

CS-W Examples

- To Be Completed
- Client (with download links)
- Server (with download links)
- Picture examples



OpenSearch

- OpenSearch is a collection of simple format for sharing of search results
- OpenSerach formats allows clients to discover and use your search engine
- Allows search engines and search clients to communicate using the common set of formats
- Was developed and created by Amazon.com and A9.com

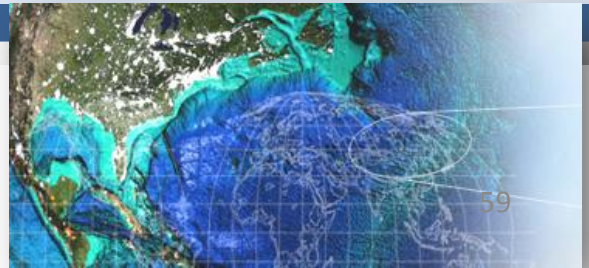
http://www.opensearch.org/Home

How To Use OpenSerach

- OpenSerach allows you to direct client to your search engine or use your search engine
  - To direct client to your web site:
    - Write a simple OpenSerach Description document to describe your search. See OpenSerach description XML document example in this link:

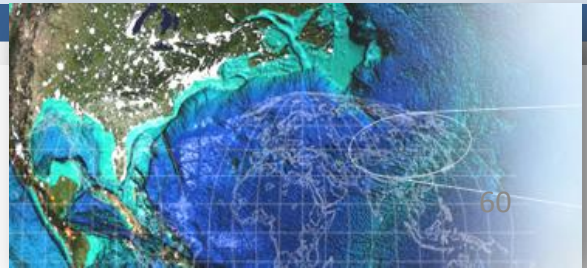
    http://www.opensearch.org/Specifications/OpenSearch/1.1#OpenSearch_description_document
- If you wish clients to use your search engine
  - Syndicate your search results by formatting them in RSS or Ato formats, augmented with OpenSerach elements. See example provided in this link:

    http://www.opensearch.org/Specifications/OpenSearch/1.1#OpenSearch_response_elements



OpenSearch Clients

- **OpenSearch search clients**
  - **Search Aggregatio Websites**
    - A9.com
    - OSfeed
    - TagJag
    - Keywop
- **Web browsers**
    - Internet Explorer 7: user instructions to install an opensearch provider- also supports OpenSearch Referrer Extension
    - Firefox 2.0 - also support OpenSearch Suggestions Extension
    - Arora - also supports OpenSearch Suggestions Extension and OpenSearch Referrer Extension
    - Google Chrome



Writing OpenSearch

- **Writing OpenSearch results with various software**
  http://www.opensearch.org/Community/OpenSearch_software
  - Alfresco
  - Drupal OpenSearch Results by Robert Douglass
  - Kwiki by Tatsuhiko Miyagawa
  - Lucene
  - Lucene by Apache.org
  - Nutch
  - OpenLink Data Spaces by OpenLink Software
  - MediaWiki by Gregory Szorc
  - Moveable Type by Alf Eaton
  - PLOS (Plone OpenSearch) (new!)
  - SearchGenerator (Ruby on Rails)
  - Wordpress by Chris Fairbanks
  - GeoNetwork opensource geospatial catalog (new!)
  - PyOpenSearch example Python Whoosh application with jQuery client

Reading OpenSearch Description Document

- **Reading OpenSearch**
  - Drupal client library by Robert Douglass
  - Drupal OpenSearch Aggregator by Steven Wittens
  - Perl library by Tatsuhiko Miyagawa and Brian Cassidy
  - PHP Hirose Masaaki
  - Python library by Ed Summers
  - ROME plugin
  - OpenLink Virtuoso by OpenLink Software
  - Apache Abdera
  - OJAX (new!)
  - LibraryFind open source metasearch application (new!)
  - qopensearch - a set of Qt classes

ESIP Federation OpenSearch

- The ESIP Federation Open Search utilize OpenSerach description XML document for:
  - Two Step Search
    1) Search for datastes and then granules within the selected dataset
    2) Space-time granule query for selected Datasete

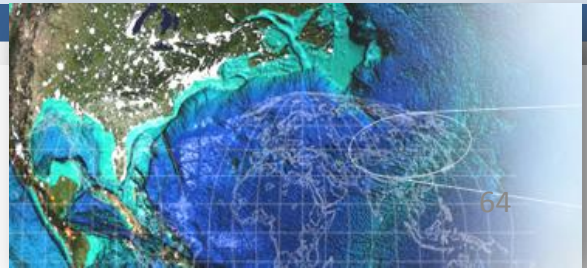
TO be Completed

http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20100003371_2010003020.pdf

Federation of Earth Science Information Partners
MAKING DATA MATTER

ESIP Federation OpenSearch

- To Be Completed

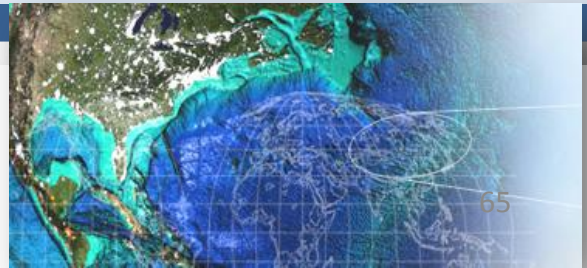


Federation of Earth Science Information Partners
MAKING DATA MATTER

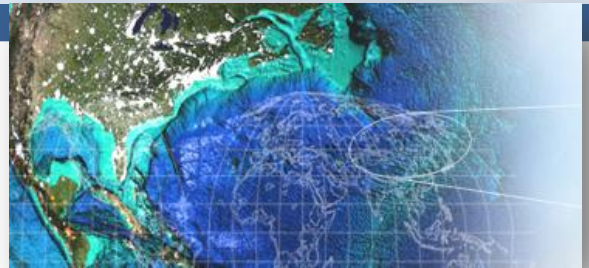
ESIP Federation OpenSearch Examples

- To Be Completed

- Client (with download links)
- Server (with download links)
- Picture examples



Usage and File Formats



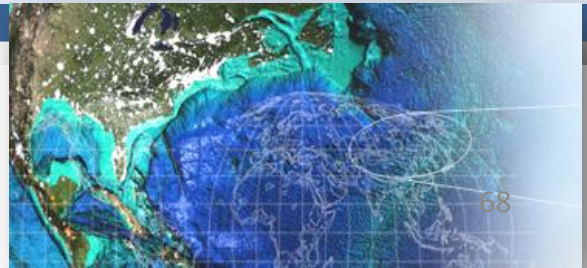
Federation of Earth Science Information Partners
MAKING DATA MATTER

Metadata and Related Standards (1/2)

- Federal Geographic Data Committee (FDGC) Metadata standards
  - FGDC Content Standard for Digital Geospatial Metadata (CSDGM) and its extension of remote sensing data to design and implement catalogue service.
- NetCDF refers to a data model for array-oriented scientific data.
- Climate and Forecast (CF) conventions for netCDF, serves as a interoperability standard.

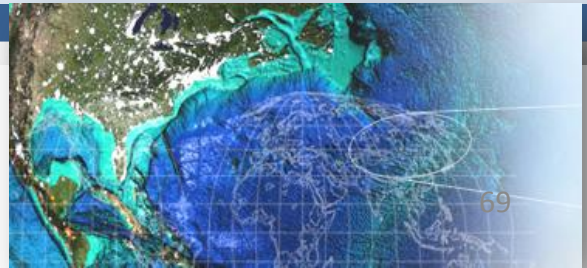
Metadata and Related Standards (2/2)

- Geospatial datasets (ISO 19115 parts I and II metadata)
- Geospatial services (ISO 19119)
- NASA ECS (EOS Core System) and remote sensing datasets
- NASA GCMD DIF (Data Interchange Format), by Tyler Stevens/GSFC (http://gcmd.nasa.gov/User/difguide/difman.html)



FGDC Metadata Standard

- To Be Completed

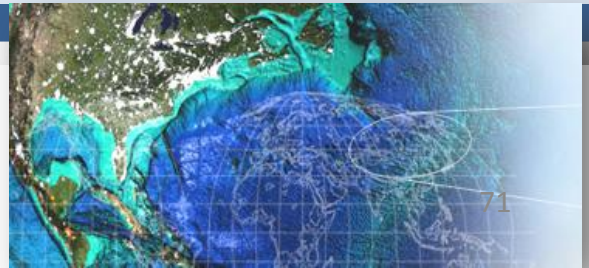


FGDC Content Standard for Digital Geospatial Metadata (CSDGM)

- "The objectives of the standard are to provide a common set of terminology and definitions for the documentation of digital geospatial data. The standard establishes the names of data elements and compound elements (groups of data elements) to be used for these purposes, the definitions of these compound elements and data elements, and information about the values that are to be provided for the data elements."
- http://www.fgdc.gov/metadata/csdgm/introduction.html

NetCDF

- "NetCDF (network Common Data Form) is a set of software libraries and machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data." http://www.unidata.ucar.edu/software/netcdf/
- Downloads are available at http://www.unidata.ucar.edu/downloads/index.jsp



Climate and Forecast (CF)

- "The conventions for climate and forecast (CF) metadata are designed to promote the processing and sharing of files created with the NetCDF API."
- "The conventions define metadata that provide a definitive description of what the data in each variable represents, and the spatial and temporal properties of the data."
- "This enables users of data from different sources to decide which quantities are comparable, and facilitates building applications with powerful extraction, regridding, and display capabilities."
- http://cf-pcmdi.llnl.gov/

Geospatial datasets (ISO 19115 part I and part II metadata)

- "ISO 19115:2003 defines the schema required for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data." http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=26020
- This is revised by ISO 19115-2:2009 (Extensions for imagery and gridded data) and ISO/NP 19115-1, which is currently in development with a target publication date around mid-2012.

Geospatial services (ISO 19119)

- "ISO 19119:2005 identifies and defines the architecture patterns for service interfaces used for geographic information, defines its relationship to the Open Systems Environment model, presents a geographic services taxonomy and a list of example geographic services placed in the services taxonomy. It also prescribes how to create a platform-neutral service specification, how to derive conformant platform-specific service specifications, and provides guidelines for the selection and specification of geographic services from both platform-neutral and platform-specific perspectives."
- http://www.iso.org/iso/catalogue_detail.htm?csnumber=39890

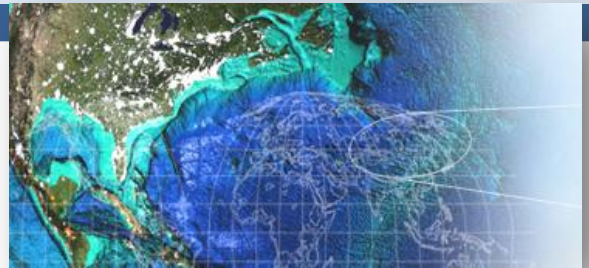
NASA GCMD DIF (Data Interchange Format)

- Originated in 1987 as the product of an Earth Science and Applications Data Systems workshop, as a step towards data system interoperability.
- As the “container” for the metadata in the Committee on Earth Observation Satellites (CEOS) International Directory Network (IDN), it does not compete with other metadata standards.
- “The DIF is used to create directory entries which describe a group of data. A DIF consists of a collection of fields which detail specific information about the data.”
- http://gcmd.nasa.gov/User/difguide/whatisadif.html

Relational Database Management System (RDBMS)

- "A **relational database management system (RDBMS)** is a database management system (DBMS) that is based on the relational model as introduced by E. F. Codd. Most popular commercial and open source databases currently in use are based on the relational database model."
- http://en.wikipedia.org/wiki/Relational_database_management_system

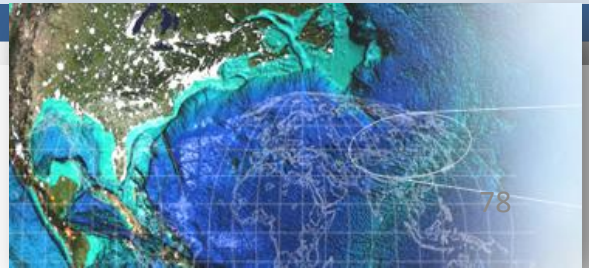
Semantic Web



Federation of Earth Science Information Partners
MAKING DATA MATTER

Semantics

- Semantics (from Greek sēmantikós) is the study of meaning. It typically focuses on the relation between signifiers, such as words, phrases, signs and symbols, and what they stand for.
- http://en.wikipedia.org/wiki/Semantics



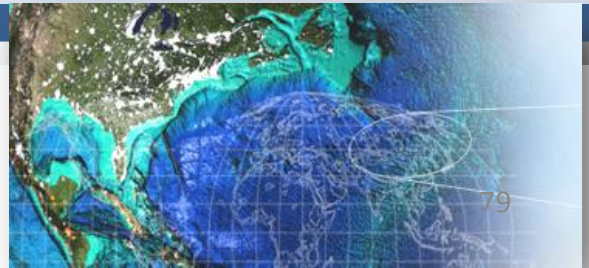
Semantic Web

- Semantic Web was thought by Tim Berners-Lee the inventor of WWW, HTTP, URLs, and HTML

http://en.wikipedia.org/wiki/Tim_Berners-Lee

- The Semantic Web is an extension of the World Wide Web through the embedding of additional semantic metadata, using semantic data modeling techniques such as Resource Description Framework (RDF) and OWL Web Ontology Language (OWL)".

http://en.wikipedia.org/wiki/Semantics



Syntactic Web vs. Semantic Web

- Syntactic Web
  - Interchange data are controlled by the application
  - Interpretation and identification of data are done by human beings
  - With the increase in volume of data and complexity, it is virtually impossible to manage these data (Information Overload)
  - Information Overload can pose a serious threat to Syntactic Web usefulness
- Semantic Web:
  - Utilizes common formats of data and combines them together from different sources
  - Provides capabilities to record how data are related to real world objects
  - Allows humans or machines to start searching one database and move on to an unending set of databases
  - Databases are not connected via wires, but by the same concepts

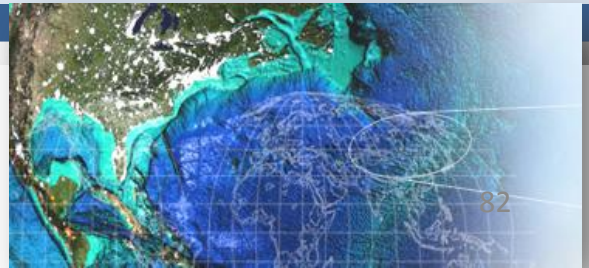
Semantic Web Example

Purchasing a new computer

- Search criteria are: 17-inch screen, 4 GB of RAM, bonus software and games, lowest available price, new, free shipping or shipping cost less than $10
- With Syntactic Web all you can do is search through different web pages and compare the conditions listed above, or use pages that compare dealers with available prices
- With a Semantic Web Agent:
  - User enters preferences listed above into computerized agent
  - Agent will initiate a complex search through invisible metadata only visible to computers.
  - Agent will display the best option, let you place an order, open your credit card payments, and mark the date of arrival

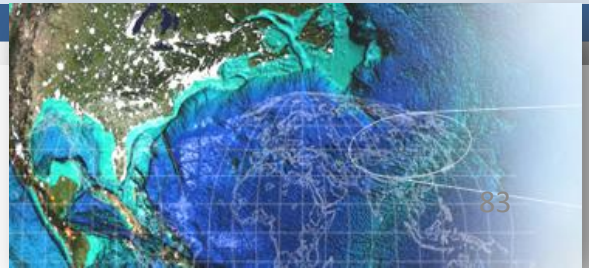
Geospatial Semantic Web (GSW)

- The idea behind GSW is to have “discovery, query, and consumption of geospatial content … based on formal semantic specification.”
- An OGC Interoperability Experiment “aims to augment WFS/FE with a semantic query capability, through the definition of an ontology for the geospatial intelligence community.”
- http://www.opengeospatial.org/projects/initiatives/gswie



GSW Interoperability

- The GSW will enable the meaning of geographic queries to be easily shared among different software systems and online services.



Federation of Earth Science Information Partners
MAKING DATA MATTER

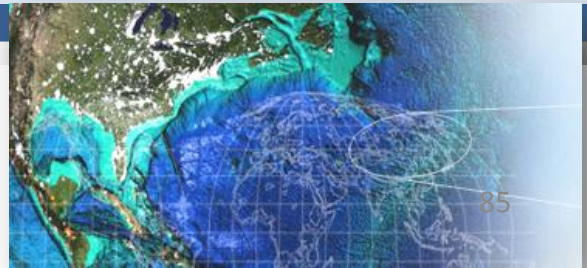
Semantic Sensor Web (SSW)

- "The **Semantic Sensor Web** (SSW) is an approach for annotating sensor data with spatial, temporal, and thematic semantic metadata. This technique builds on current standardization efforts within the Open Geospatial Consortium (OGC) Sensor Web Enablement (SWE) and extends them with Semantic Web technologies to provide enhanced descriptions and access to sensor data"

http://en.wikipedia.org/wiki/Semantic_Sensor_Web

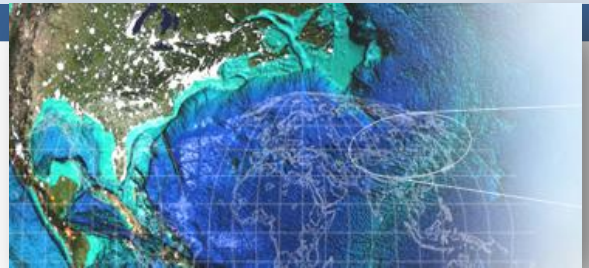
SSW Interoperability

- To Be Completed



Federation of Earth Science Information Partners
MAKING DATA MATTER

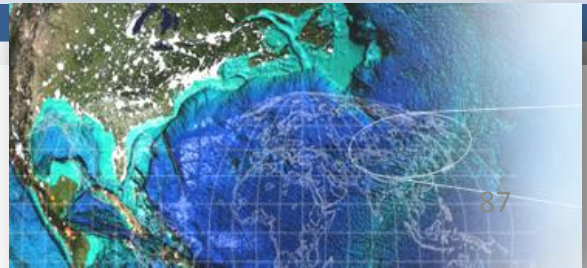
Case Studies



Federation of Earth Science Information Partners
MAKING DATA MATTER

Case Studies

- Case studies illustrate the benefits of using interoperability standards.
- The following case studies are presented:
  - MODAPS WCS Web Services
  - The NOAA-led Integrated Ocean Observing System (IOOS)
  - NASA's sensor webs

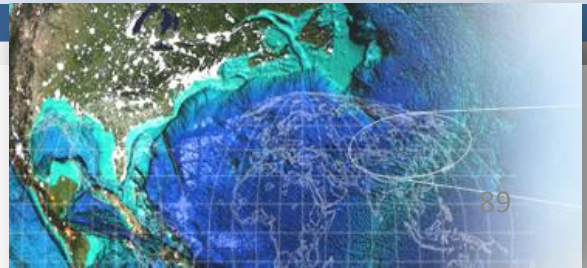


MODAPS WCS Web Services (1/3)

- Provides a SOA machine-to-machine API to level 1 and atmospheric archived MODIS (Moderate Resolution Imaging Spectroradiometer) data
- Utilizes standard Web Service interfaces through both Simple Object Access Protocol (SOAP) and Representational State Transfer (REST) protocols
- Allows OpenSearch and the Open Geospatial Consortium's Web Coverage Service (WCS) interfaces

MODAPS WCS Web Services (2/3)

- Provide synchronous web services that allows users to utilize WCS capabilities to perform:
  - OpenSearch and Open Geospatial Consortium
  - Gridded search
  - Post processing



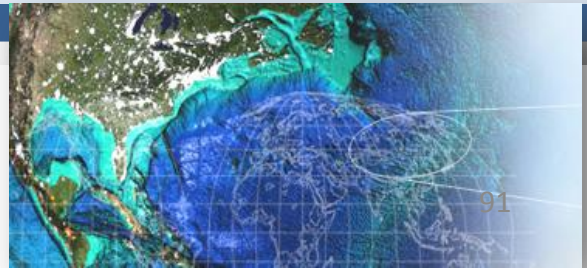
Federation of Earth Science Information Partners
MAKING DATA MATTER

MODAPS WCS Web Services (3/3)

- MODAPS WCS Web Services Implementation Issues, Pros, and Cons
  - WCS Issues and Concerns
    - Implementation issues using AXIS2
    - MODAPS Data volume and WCS capability
    - WCS Schema Validation and Java Binding dilemma
  - Our Approach
    - Considering other Web Services Framework
    - Using automated schema binding technologies

MODAPS WCS Web Services – Issues and Concerns

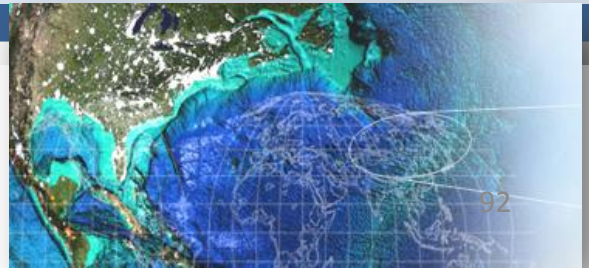
- SOA and AXIS2
  - Security
  - AXIS2 Software Support
  - Contract last vs. Contract first



Federation of Earth Science Information Partners
MAKING DATA MATTER

MODAPS WCS Web Services – WCS Schema Issues

- OGC WCS Schema
  - WCS Schema fails to support MODAPS's large number of products
  - XMLSpy fails to validate OGC WCS Schema version 1.0.0
  - Castor Binding tools fail to create Java classes to support this Schema
  - WCS Schema many based objects that are being restricted
    - This schema design cause failures within Castor Binding



MODAPS WCS Web Services – Lessons Learned

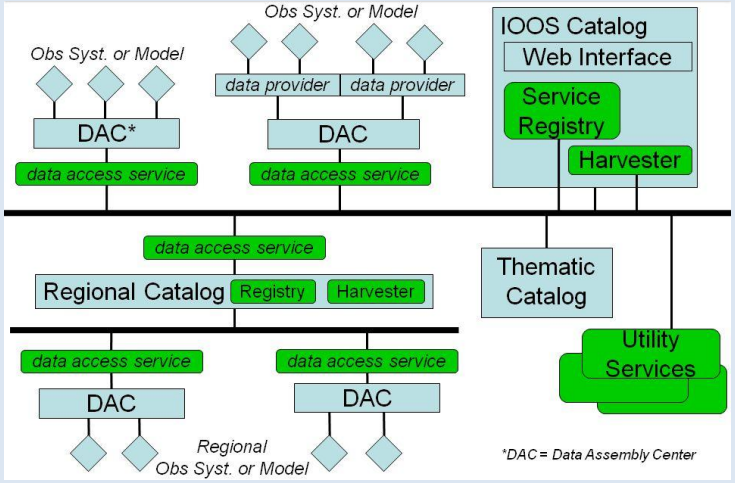
- Use experience with other technologies such as Spring Web Services instead of AXIS2 to eliminate:
  - Dependencies to AXIS2 OMElement
  - AXIS2 Security issues
  - AXIS2 lack of software and security support
- Use XML binding COTS products such as CASTOR to eliminate manual implementation of Java API to handle OGC WCS Schema.
- Difficulty with validating WCS Schema, modifying the schema to be readable by CASTOR Binding tool and the same time protect the integrity of WCS Schema standards.

The NOAA-led Integrated Ocean Observing System (IOOS)

- "The Integrated Ocean Observing System (IOOS) is a coordinated network of people and technology that work together to generate and disseminate continuous data on our coastal waters, Great Lakes, and oceans."
- "By collecting and bringing data together in a way that ensures the information can be used with other data sets, IOOS will make a broader suite of data available to scientists, allowing them to develop a more complete characterization of our oceans and coasts. "
- "IOOS is a major shift in our approach to ocean observing, drawing together many networks of disparate, Federal and non-Federal observing systems to produce data, information, and products at the scales needed to support decision making."

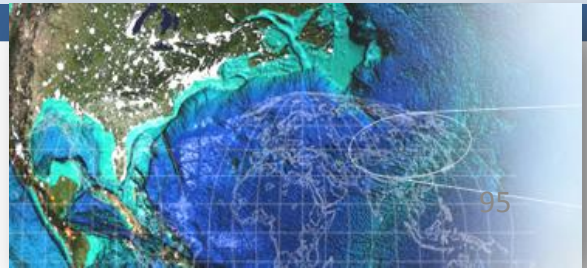
http://www.ioos.gov/about/basics.html

About IOOS



This figure shows an IOOS federated, service-oriented architecture.

- The Integrated Ocean Observing System (IOOS) provides information about open oceans and US coastal waters and Great Lakes to scientists, managers, businesses, governments, and the public in order to support research, to inform decision-making, and to enable new applications and derived products beyond the original intent of the data gathering.
- Programmatically, IOOS is a component of the Global Ocean Observing System (GOOS) and the Global Earth Observing System of Systems (GEOSS). The US National Oceanic and Atmospheric Administration (NOAA) has been assigned the role of lead federal agency in this endeavor.
- Technically, IOOS includes or interfaces with existing observing systems, data providers, and archives, and IOOS collaborates in developing additional capabilities in observations, data management and data use.
- The IOOS uses mostly OGC Sensor Observation Service (SOS) and Unidata's Data Access Protocol (DAP), and to a somewhat lesser extent (so far) OGC Web Coverage Service (WCS) and Web Map Service (WMS) Technical guidance document: http://ioos.gov/library/dmac_implementation_2010.pdf



IOOS Data Access

- "The NOAA IOOS program initiated development of a Data Integration Framework (DIF) to improve management and delivery of an initial subset of ocean observations."
- "The following services are the first to be established by the NOAA IOOS program and its partners to provide access to data."
- National Data Buoy Center (NDBC) Sensor Observation Service (SOS) (http://sdf.ndbc.noaa.gov/sos/)
- CO-OPS SOS (http://opendap.co-ops.nos.noaa.gov/ioos-dif-sos/)
- NDBC THREDDS Data Server (http://sdf.ndbc.noaa.gov/wcs/)
- SECOORA SOS (http://neptune.baruch.sc.edu/xenia/sos/difSOS.html)

http://www.ioos.gov/dif/

Sample NDBC SOS Output

- GetObservation for WaterLevel (sea_floor_depth_below_sea_surface)
- Single point, Station 46403, observation for a specific time, CSV format
- http://sdf.ndbc.noaa.gov/sos/server.php?request=GetObservation&service=SOS&offering=urn:ioos:station:wmo:46403&observedproperty=sea_floor_depth_below_sea_surface&responseformat=text/csv&eventtime=2008-07-17T00:00Z

```
station_id,sensor_id,"latitude
   (degree)","longitude
   (degree)",date_time,"sea_flo
   or_depth_below_sea_surface
   (m)","averaging interval
   (s)"
urn:ioos:station:wmo:46403,urn
   :ioos:sensor:wmo:46403::tsun
   ameter0,52.65,-156.94,2008-
   07-17T00:00:00Z,4509.488,900
```

Sample CO-OPS SOS Output

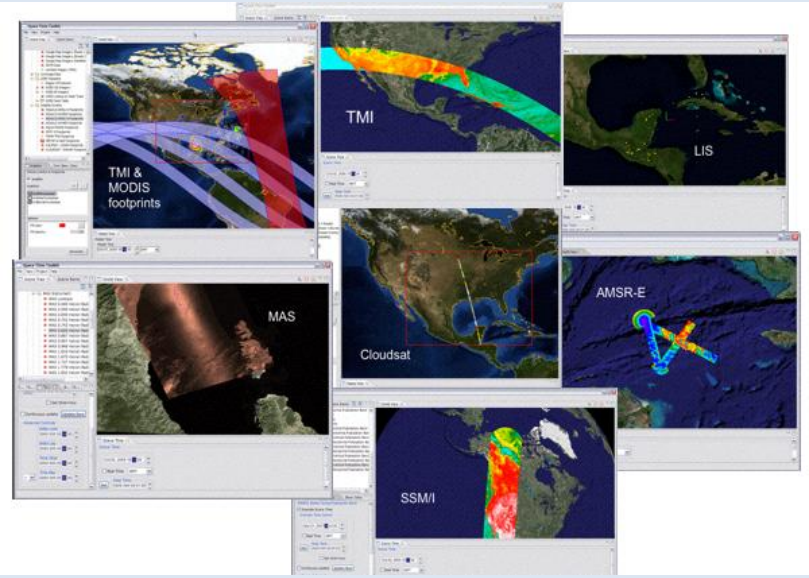
- Currents Data (GetObservation Service)
- Profile Bin Data (Latest)
- http://opendap.co-ops.nos.noaa.gov/ioos-dif-sos/get/currents/currentsprofile.jsp
- Station db0301
- Result format text/xml

```
…
 <ioos:Count name="Station1T1NumberOfBinObservations">40</ioos:Count>
 <ioos:ValueArray name="Station1T1ProfileObservations" gml:id="S1T1PO">
   <gml:valueComponents>
     <ioos:CompositeValue name="Station1T1Bin1Obs" gml:id="S1T1B1" processDef="#Station1Sensor1Info">
       <gml:valueComponents>
         <ioos:Quantity name="CurrentDirection" uom="deg">199.0</ioos:Quantity>
         <ioos:Quantity name="CurrentSpeed" uom="cm/s">39.1</ioos:Quantity>
       </gml:valueComponents>
     </ioos:CompositeValue>
…
```

IOOS Software

- The IOOS also provides some software “on an as-is basis with no support or warranty”, including
- Server Code
  - NDBC SOS software version 1
  - GCOOS SOS software beta version 0.6.1
  - THREDDS Data Server (TDS)
- Format Converters
  - IOOS SOS to CSV v0.6.1
  - netCDF to BUFR converter beta v0.1
- http://www.ioos.gov/dif/

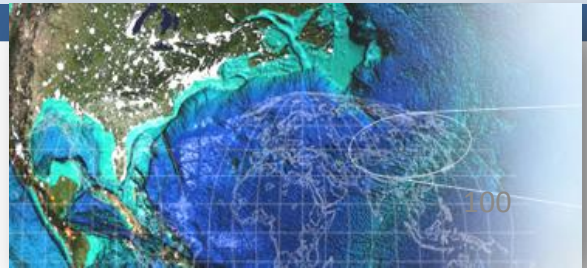
NASA's Sensor Webs (1/4)



- A number of NASA projects have adopted the OGC SWE suite of standards.
- Collaboration between JPL and GSFC with the Earth Observing 1 (EO-1) satellite is an important part.
- EO-1 was used in a wildfire sensor web scenario, and GSFC with partners prototyped a transformation to an SWE framework using GeoBliki (see http://eo1.geobliki.com)

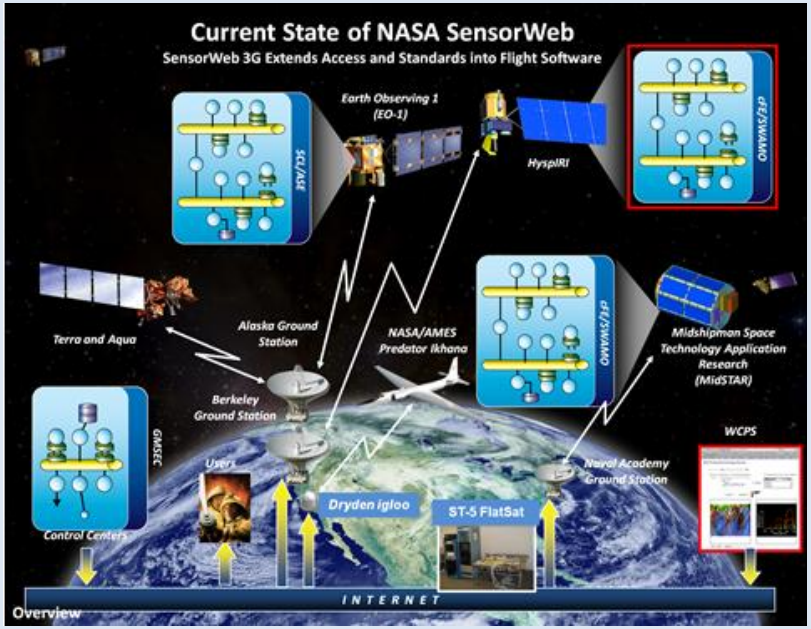
All of these satellite and airborne sensors are, at least some of the time, using SensorML for geolocation and other purposes. See NASA's JPL and GSFC Sensorweb/EO-1 pages, http://sensorweb.jpl.nasa.gov/ and http://eo1.gsfc.nasa.gov/

http://eo1.gsfc.nasa.gov/new/validationreport/Technology/Documents/Tech.Val.Report/SWE%20article%20for%20Dec07%20Sensors%2020071024.doc



Federation of Earth Science Information Partners
MAKING DATA MATTER

NASA's Sensor Webs (2/4)



- "The main objective of the sensorweb activity is to create an interoperable environment for a diverse set of satellite sensors via the use of software and the Internet … to better understand physical phenomena … [and] it facilitates science investigation."
- "The end goal is to make discovery and access to sensors as easy as finding and using websites on the Internet."

Figure and information from http://eo1.gsfc.nasa.gov/new/sensorWebExp/index.html

NASA's Sensor Webs (3/4)

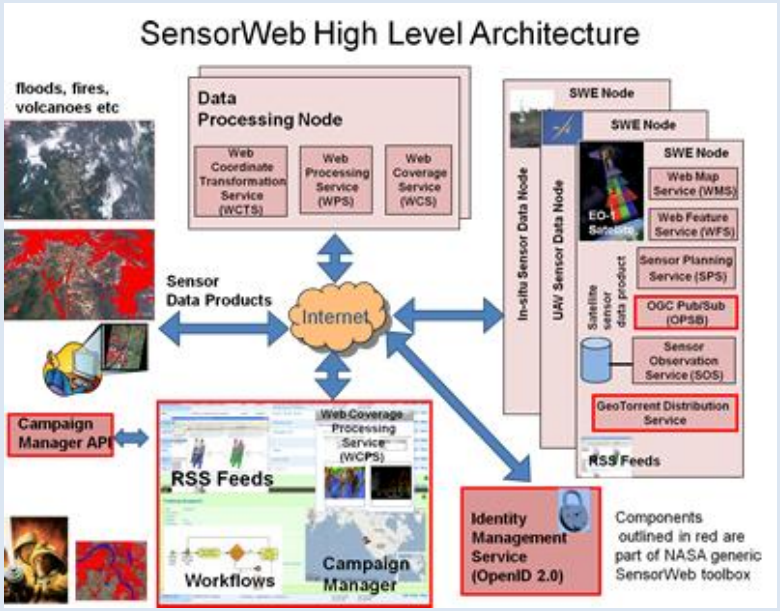
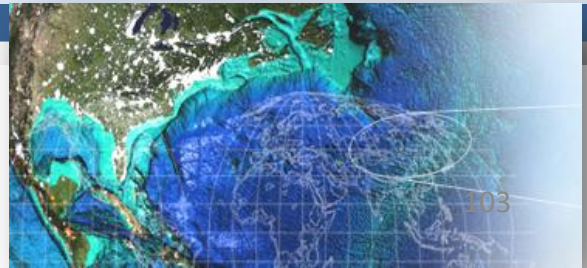


Figure and information from http://eo1.gsfc.nasa.gov/new/sensorWebExp/Components.html

- The basic components are:
  - Reference architecture
  - A set of web services
  - Cross-domain generic scripting language
  - Campaign Manager 1.0
  - Campaign Manager 1.0 API
  - Identity Management Service
  - OGC Publish/Subscribe
  - GeoTorrent
- The five components outlined in red in the figure at left were created by NASA to augment the international standards.
- These are NASA intellectual property, but will be available as open source software.

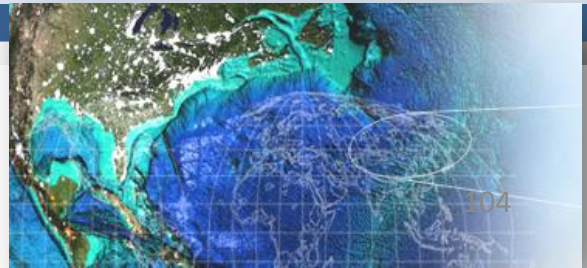
NASA's Sensor Webs (4/4)

- Nearly 100 papers, over 30 presentations, and 16 articles on NASA's sensor web efforts are listed at http://eo1.gsfc.nasa.gov/new/sensorWebExp/Papers.html
- Subject areas include
  - An interoperable sensor architecture
  - Autonomous sciencecraft experiments
  - On-board diagnostics tools
  - Adaptive phased array ground antennas



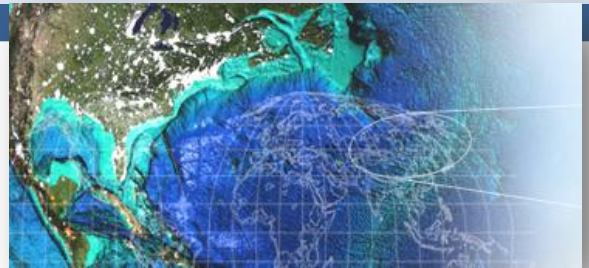
Additional Resources

- A document on Practical Data Interoperability for Earth Sciences is available at http://www.esdswg.org/techinfusion/downloads/pdies



Federation of Earth Science Information Partners
MAKING DATA MATTER

Computing Platforms



Federation of Earth Science Information Partners
MAKING DATA MATTER

Computing Platforms

To Be Completed

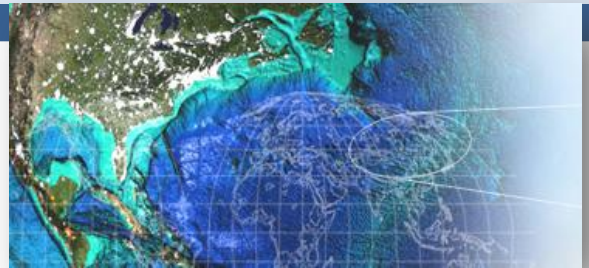
Note From ESIP Reviewer:

Please explicitly address information about computing platforms that may be needed to implement and use standards based Earth Science Products instead of mentioning "hardware/software"

Rahul Comment:

You could address here something specific for example Service Oriented Architecture listing both SOAP and Rest based services as software based approach suited for easy interoperability.

References



Federation of Earth Science Information Partners
MAKING DATA MATTER

References

citations listed in order of referenced in this presentation

New World Encyclopedia, http://www.newworldencyclopedia.org/entry/Interoperability, October 10, 2008

WCS Benefits http://docs.geoserver.org/latest/en/user/services/wcs/reference.html, Open Geospatial Consortium, copyright 2009 GeoServer.

CSW Interface (2.02) http://eie.cos.gmu.edu/CSWClient/interface202.jsp, George Mason University, 2009, CISC, George Mason University 2009.

Sensor ML - http://www.opengeospatial.org/standards/sensorml Open Geospatial Consortium, Inc. 1994

Botts, Mike, " Sensor Model Language: Moving Sensor Data on to the Internet", April 1, 2003, Questex Media Group LLC, http://www.sensorsmag.com/networking-communications/a-sensor-model-language-moving-sensor-data-internet-967

NASA GCMD DIF (Directory Interchange Format (DIF) Writers Guide, 2010. Global Change Master Directory, National Aeronautics and Space Administration. [http://gcmd.nasa.gov/User/difguide/] (http://gcmd.nasa.gov/User/difguide/difman.html)

OPeNDAP - http://opendap.org/ , 2008, OPeNDAP, Inc.

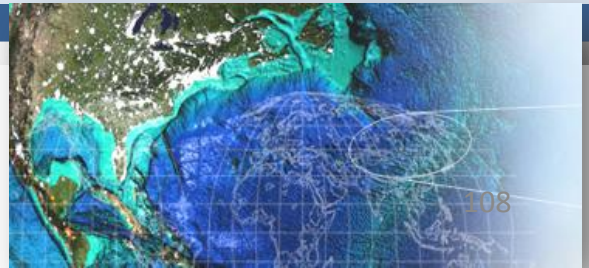
OPeNDAP at GSFC http://disc.gsfc.nasa.gov/services/opendap/ , NASA's Science Mission Directorate (SMD), archived and distributed by Goddard Earth Sciences (GES) Data and Information and Services Center (DISC).

http://disc.sci.gsfc.nasa.gov/gesNews/opendap_AIRS_data_access, NASA's Science Mission Directorate (SMD), archived and distributed by Goddard Earth Sciences (GES) Data and Information and Services Center (DISC).

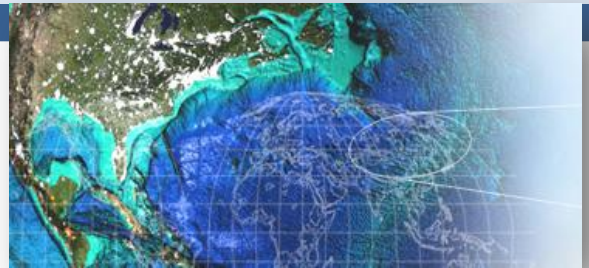
Codd, E.F., "A Relational Model of Data for Large Shared Data Banks", Communications of the ACM, Volume 13, Number 6, June 1970

Stock, Kristin, "OGC Catalogue Services – OWL Application Profile of CSW", July 2009, project document OGC 09-010 version 0.3.0, Open Geospatial Consortium, Inc.

deLaBeaujardiere, Jeff, "Building the IOOS Data Management Subsystem", Marine Technical Society Journal, IOOS Special Journal (Submitted 2010)



Thank you!



Federation of Earth Science Information Partners
MAKING DATA MATTER

Backup Slides

Federation of Earth Science Information Partners
MAKING DATA MATTER

GSFC DAAC (1/4)

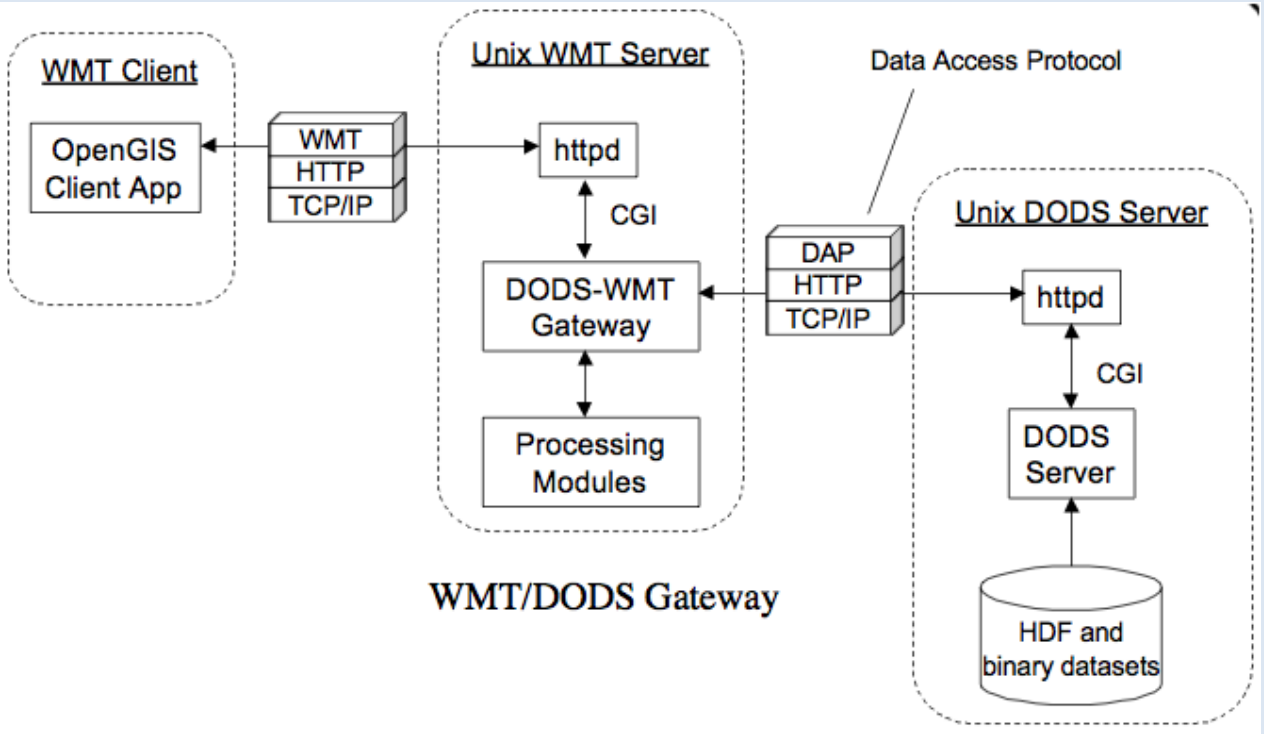
- The GSFC DAAC is the Goddard Space Flight Center's Distributed Active Archive Center
- This case study shows the interoperability between Digital Earth and DAAC's Distributed Oceanographic Data System (DODS).
- It offers a variety of data sets requiring multiple rendering routines.

Federation of Earth Science Information Partners
MAKING DATA MATTER

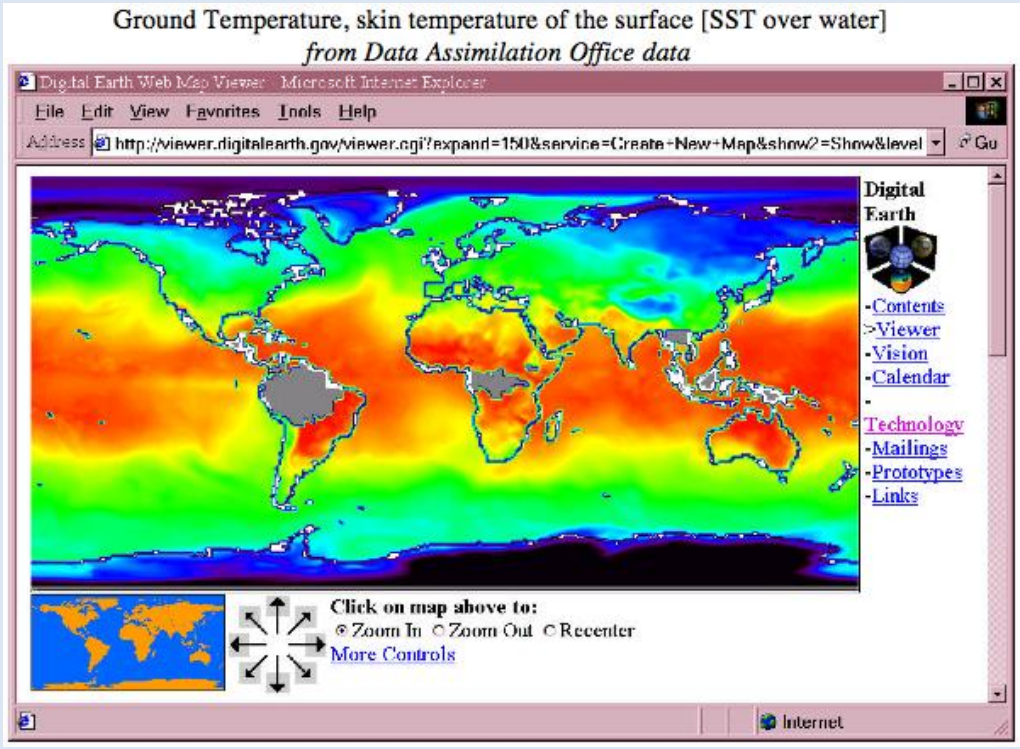
GSFC DAAC (2/4)

- Users access data over the network in a consistent manner.
- The DAAC provides access to multiple DODS servers via one gateway.
- It uses Web Mapping Testbed (WMT), Data Access Protocol (DAP), OPeNDAP Grid Protocol, OpenGIS, and HTTP.

GSFC DAAC (3/4)


WMT Client
OpenGIS Client App
WMT
HTTP
TCP/IP
Unix WMT Server
httpd
CGI
DODS-WMT Gateway
Processing Modules
Data Access Protocol
DAP
HTTP
TCP/IP
Unix DODS Server
httpd
CGI
DODS Server
HDF and binary datasets
WMT/DODS Gateway

GSFC DAAC (4/4)


Ground Temperature, skin temperature of the surface [SST over water]
from Data Assimilation Office data
Digital Earth Web Map Viewer - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://viewer.digitalearth.gov/viewer.cgi?expand=150&service=Create+New+Map&show2=Show&level
Go
Digital Earth
-Contents
>Viewer
-Vision
-Calendar
-Technology
-Mailings
-Prototypes
-Links
Click on map above to:
Zoom In Zoom Out Recenter
More Controls
Internet